Cross-Industry Analysis of Russian Enterprise Performance: Do Concentration and Diversification Matter?

Zyuzin, Andrey

Экономическая политика. 2021. Т. 16. № 4. С. 42–83 DOI: 10.18288/1994-5124-2021-4-42-83

Cross-Industry Analysis of Russian Enterprise Performance: Do Concentration and Diversification Matter?

Alexander ZYUZIN

Alexander V. Zyuzin —
PhD student, National Research University Higher School of Economics
(11, Pokrovskiy bul., Moscow, 109028, Russian Federation).
E-mail: azyuzin@hse.ru

Аннотация

This paper studies and quantifies the influence of industry localization and regional economic diversification on firm performance in different industries. The main idea is that industry localization and regional economic diversification improve enterprise performance and the influence might be found for most industries. In this paper, the Ellison-Glaeser index is applied to measure localization effects and the Herfindahl-Hirschman index is used to measure diversification. The dataset consists of 650,000 observations and approximates the full set of Russian real sector commercial companies in 2017. All companies were aggregated into eight groups by respective industry. Firm performance is measured via sales margin and net profit margin. Since the performance is highly dispersed, the regression analysis includes both OLS and quantile regression (QR) models for each group. It was found that the Mining industry had been affected neither by localization nor by diversification. Localization effects are significant and positive for Agriculture, Fishing, and Forestry; Wholesale, Retail and Food Services; Manufacturing; Transport; and IT, Broadcasting, and Telecommunication. The increase in profitability for them ranges from 1% to 4% per 0.1 change in the Ellison-Glaeser index. Localization is harmful for Construction, and Services and Minor Industry companies (7-fold drop in sales margin for Construction). Diversification is significant and decreases the sales margin of all the companies, and the effect ranges from 1.01% to 1.22% per 0.1 change in the Herfindahl-Hirschman index. These findings hold despite the choice of analysis tool (QR versus OLS); however, the study of different quantiles sheds light on the effects specific for some industries.

Keywords: regional diversification, industry localization, sales margin, net profit margin, Russian regions, enterprise performance, real sector industries, missing values estimation.
JEL: C21, R11, R12, R15.

Acknowledgements

The author sincerely thanks Vladimir Salnikov and the Center for Macroeconomic Analysis and Short-Term Forecasting (CMASF) for providing access to the ample dataset used here. The author is also grateful to HSE professors—Olga Demidova, Tatyana Dolgopyatova and Hartmut Lehmann—for their feedback and support.

Introduction

This paper studies the effects of regional economic diversification and industry localization on Russian firms' performance. Industry localization (or *"localization"*) is measured with the Ellison-Glaeser index (*EG*) [Ellison, Glaeser, 1997], and regional economic diversification (or *"diversification"*) is measured with the Herfindahl-Hirschman index (*HHI*). The dataset used for the research includes information about the financial statements of 650,000 Russian real sector private companies in the year 2017. These data closely approximate the general population for the year. One of the main methodological issues is separate analysis for the different industries as the nature of economic activity is heterogeneous and this reflects in implied firms' performance. Moreover, the study distinguishes well and poorly performing firms and measures the impact of localization and diversification for them by applying quantile regression (QR) approaches.

The paper extends the related research [Davidson, Mariev, 2015; Zyuzin et al., 2020], but pursues a different goal—to study industry localization and regional economic diversification effects on productivity of firms that belong to different industries. The firms are divided into eight groups by their OKVED[1] code and analyzed separately. The author is not aware of any paper where such an analysis had been held for enterprises discriminated by the industry factor. Although in their recent paper [Zyuzin et al., 2020] the researchers applied discrimination by business scale, all the industries were considered in a single sample. The firm scale is still controlled here via dummy implementation.

The present paper brings additional contribution to the existing literature as it captures the changes in localization and diversification effects for high and low margin firms via estimation of quantile regressions. Separate consideration of polar performance firms is justified as it allows one to distinguish the low base effect and sheds light on whether there is any relation between the power of studied effects and firms' implied performance rate. If ex-ante expectations about the diminishing value, added by industry localization and regional economy diversification into the firms' performance along with the rise of implied firms' margin, are not justified, then the firms are indeed homogeneously influenced by the studied effects.

In the study, localization and diversification are assumed to be externalities that arise from agglomerative economic mechanisms. ***Localization*** is defined closely to the classical *agglomeration economies* (and is called *"cluster"* as well): "a cluster of firms that belong to the same industry and are located in the same region in relatively greater numbers

[1] Russian national classifier of economic activities.

than in other regions." ***Diversification*** as a term is also defined here closely to its classical analog *"diversity economies"*: "a region is called *diversified* if all the industries are presented in the region in equal shares, and *specialized* if there are industries that employ significantly greater shares of labor than others."

The classical division between urbanization economies and localization economies [Cohen, Morrison Paul, 2009; Jacobs, 1961] cannot be directly applied in this paper—however, it is close to the introduced diversification and localization terms. The difference is that this study is based on regional-level data rather than city-level data which are typically used for urbanization economies studies. This approach is justified, primarily, since Russia is a federal state and each region has wide credentials in the social and economic spheres (either direct or indirect), including regulation of local taxation rates, wage coefficients for officials or pension sizes. Moreover, its regions are also highly dispersed in terms of their natural characteristics (such as climate, resource field or border position), which might be significant for some industries. In addition, it was found that regions influence each other through spatial externalities the same way as minor administrative units [Kolomak, 2010]. Finally, some controlling variables used in this study are only available at the regional level of aggregation.

The main hypothesis tested in this research is composite and consists of two parts that are formulated as follows:

a. "Localization and diversification effects simultaneously exist for Russian enterprises, and their scale and direction of influence on firms' performance are different depending on the industry as well as regional and internal companies' characteristics";
b. "Depending on the industry, the influence of localization and diversification effects differs for well and poorly performing companies, and some non-linear relations could be verified."

There is a lot of evidence that agglomeration effects influence companies in a number of ways, and those findings are described in the empirical review section. In the next section follows an overview of the theoretical background, developed in the past century, where it has been suggested that localization should boost companies' performance [Marshall, 1920; Rosenthal, Strange, 2004] while diversification effects might be ambiguous [Jacobs, 1961; Quigley, 1998].

1. Review of Agglomeration Theory Development

The earliest hypotheses about the existence of agglomeration effects (or agglomeration externalities) were formulated at the beginning of the

20th century and are usually associated with Alfred Marshall's *Principles of Economics* [Marshall, 1920]. Marshall's theoretical framework for localization economies is simple and is based on the existence of positive externality effects that come from three channels: *labor pooling*, *knowledge spillovers* and *consumer-supplier chain enhancement*. The theory was developed and enriched by a number of works, but the main ideas remained unchanged [Duranton, Puga, 2004; Rosenthal, Strange, 2004].

(a) Labor pooling results in the opportunity for the firms to hire employees as lower costs. The existence of this channel has been shown in a number of works [Baumgardner, 1988; Cohen, Morrison Paul, 2009; Nakamura, 1985].

(b) Knowledge spillover effects boost technical progress due to enhanced and intense information and knowledge exchange, and cooperation in knowledge production [Audretsch, 1998; Audretsch, Feldman, 1996; Morrison Paul, Siegel, 1998; Wheaton, Lewis, 2002].

(c) Better access to both suppliers and consumers results in a decrease of logistics costs and economic disintegration, and thus the firms also gain benefits from this externality [Cohen, Morrison Paul, 2009; Holmes, 1999].

Diversification effects stem from less obvious and indirect externalities: benefits from social protection and general faster regional economic growth [Jacobs, 1961].

Along with the progress in computer science and econometric methods, an increasing scope of empirical papers appeared where evidence for agglomeration economies was found. Modern papers on the topic usually provide empirical testing of existing theories rather than search for new channels that connect agglomeration and firms' behavior. The next section reviews empirical papers closest in terms of methodology and research strategy to the present one.

2. Review of Empirical Papers

One of the early works that studied the connection between vertical disintegration (measured close to the sales margin used in this paper) and localization was introduced in 1999 [Holmes, 1999]. To measure the degree of localization, the author used a purchased input intensity coefficient. The research was based on cross-section US manufacturing company data that included 368,000 observations for the year 1987. With a weight OLS model applied, a positive relation was found between localization and disintegration.

A similar research was held for Chinese manufacturing companies [Lu et al., 2012]. The authors studied 241,000 firms belonging to the manufacturing industry (214,000 in the main group and 27,000 in the control group) and constructed a panel for the period from 1998 to

2005. The question of interest was how agglomeration influences firms' market markup. A major point in the study is the measurement of the markup, which the authors define for each firm, following the well-known methodology [Loecker, Warzynski, 2012], as the ratio between price and the firm's marginal costs. The degree of agglomeration was measured using EG [Ellison, Glaeser, 1997]. The researchers show that agglomeration decreases the markup.

Another important paper in the literature is the one dedicated to Russian trade, catering and manufacturing companies [Davidson, Mariev, 2015]. The authors tested the hypothesis that the firms' revenue depends on industry localization and city-level economic diversity. Localization was measured as the logarithm of difference between the total revenue that firms of an industry gain in a certain region for a given year and a single i^{th} firm's revenue from the same industry and for the same year. Thus, in this case, pure agglomeration economy was considered. More than 7,000 companies were observed to collect a panel for the period from 2002 to 2008. The OLS estimation shows that city-level diversity increases revenue and that the effects from localization are non-linear and have a U-shaped form.

Recent studies also find strong evidence in support of the existence of agglomeration effects for Russian companies. Using the data on more than 500,000 companies, it was shown that the larger is the agglomeration size the better is the firms' average labor productivity [Lavrinenko et al., 2019].

For Russia, there are also well-known studies dedicated to agglomeration dynamics in general. After the collapse of the USSR and the establishment of market economy in Russia, the agglomeration processes intensified and industrial activity concentrated around big cities [Kolomak, 2015b; Mikhailova, 2016].

There are only a few papers that link regional economic diversification to any kind of firm-related performance measures [Davidson, Mariev, 2015; Lavrinenko et al., 2019; Lu et al., 2012]. The key theoretical ideas that support diversification economies were formulated in Jane Jacobs' books [Jacobs, 1961]. One of them showed that the yield growth of diversified regions is faster than their population growth, thus boosting the whole country's economy [Quigley, 1998]. This effect was also demonstrated applicable for the Russian economy [Kolomak, 2015a].

There are papers that study indirect agglomeration effects such as knowledge spillovers that include innovation boost effects as a proxy for knowledge production. Innovation activity is usually measured as the number of registered patents or new products introduced to the market per defined period. This question was studied in the works [Baptista, Swann, 1998; Hervas-Oliver et al., 2018]. Using a panel data set of 248 British manufacturing companies from 1975 to 1985 and applying the Poisson regression model, Rui Baptista and Gavin Swann found that localization enhances innovative activity. Jose-Luis Hervas-Oliver and colleagues got the same result for Spanish enterprises (6,700 companies cross-section for 2001, where private companies of all industries were included).

Some researchers go beyond classical channels of agglomeration to firm effects and check whether agglomeration may cause different types of externalities. For example, for Swedish exporting companies it was found that both localization and urbanization increases gross exports[2] [Malmberg et al., 2000].

This paper is in line with the empirical works discussed above, but differs from and supplements them in a number of ways. The major contribution is the comparison of localization and diversification effects for different industries provided here. Next, the methodology is compiled to discriminate firms by their implied performance and verify the change in the localization and diversification effects for high and low margin companies.

The modeling technique also differs slightly from the papers above. The enterprise performance here is measured straightforwardly through sales margin and net profit margin, in contrast to Natalia Davidson and Oleg Mariev who had taken the revenue logarithm as a dependent variable [Davidson, Mariev, 2015]. The present analysis was conducted not for cities but for larger spatial objects, namely regions, which allows all Russian regions to be covered in the paper. The inputs for the model (EG and HHI) were estimated endogenously. This is robust since the dataset is de facto the whole population of Russian firms from all the industries. The missing values contained in *Employees* (the key variable for calculating EG and HHI) were not cut but estimated instead.

3. The Data

In this section, the data sample is briefly described. Each observation is one real sector enterprise, but the firm is characterized not only through its own financial indicators but also externally from the perspective of the regional environment where the firm operates.

The sample contains 650,000 firms[3]. Initially the raw data set included around 2.5 million observations, but it shrank due to the exclusion of financial industry firms, observations with a priori incorrect or inapplicable data (unrealistically high values for revenue, assets etc., incorrect signs before financial indicators were the exclusion triggers, or observations with too many missing or zero values in their financial statements), non-commercial companies, and government-, state- or

[2] Their analysis was conducted on a sample of 10,000 Swedish export companies in the year 1994 (cross-section data). The result was captured with a simple OLS regression analysis.

[3] Sources of the firm-level data: Federal State Statistics Service, Federal Tax Service, and Center for Macroeconomic Analysis and Short-Term Forecasting (CMASF) database; source of the regional data: Federal State Statistics Service.

municipally owned enterprises[4]. Nevertheless, the dataset is not only representative, but appears to be a general population of privately owned companies for the year 2017.

Among the available firm-level data there is a full list of P&L and balance sheet indicators, including such important firm characteristics as *assets, capital, number of employees, revenue, cost of sales* and *debt*. Regional data[5] describes the external working environment of each enterprise and contains information about *unemployment, gross investments, regional R&D expenses, average monthly wages PPP, crime rate, weight of unprofitable companies* and *firm birth rate*.

The *"employees"* variable is the crucial one for the present research and deserves a separate discussion. It is provided by the Russian Federal Tax Service and is, methodologically, obtained by averaging the annual employment of each firm. This variable is the only one in the dataset that contains missing values (14% of all observations). As it is used to calculate localization and diversification, the missing values were estimated to avoid distortions in the indices' order and value. The values could not be removed from the consideration as the nature of missing values is unknown. This means that their distribution could be non-uniform among regions or industries, which might bring serious bias in EG or HHI (or both).

The estimation of *"employees"* was performed via the OLS and the multiple imputation (*MICE*) predictive mean matching methods separately to ensure robustness. In both cases the predictions were based on the available data about *assets, capital, revenue*, and *scale dummy*. Then the estimation results were compared to each other as well as possible distortions in the indices' values that arise from switching from one method to another. The MAPE[6] criterion measured an error among estimated sets generated by the OLS and the MICE pmm models. An error did not exceed 15% whatever initial parameters were set in the multiple imputation function. After that EG and HHI were calculated by using each of the imputed datasets and the results were compared. The high value of the Spearman correlation coefficient (0.96) guarantees preservation in the order of the values while switching between OLS and MICE pmm generated samples. The absolute values of the indices also turned out to be close (Pearson correlation among the correspond-

[4] Non-commercial or state-owned companies were excluded from the analysis as their goal might differ from classical profit maximizing agents. For example, such companies often invest in socially important or infrastructure projects, such as building hospitals or schools.

[5] Note that only 82 Russian regions are considered in this paper. The reason for this is that Nenets Autonomous Okrug was included into Arkhangelsk region, whereas Khanty-Mansi Autonomous Okrug and Yamalo-Nenets Autonomous Okrug were both included into Tyumen region.

[6] $MAPE = \frac{1}{n}\sum_{t=1}^{N}\left|\frac{A_t - F_t}{A_t}\right|$ — mean absolute percentage error, where A_t is the actual value (one got by the OLS method in our case) and F_t is an alternative forecasted value (one got by the MICE pmm). MAPE was calculated for two artificially estimated samples and does not exceed 15%, which means that both methods provide estimations close to each other.

ing indices was not exceeding 0.93). Since the estimations appeared to be close to each other, only one sample generated by the OLS method was chosen for the consecutive work.

4. Methodology

This section is dedicated to the model specification and consists of a number of issues: the aggregation of firms by industry groups, the description of each group (descriptive statistics), the calculation of localization and diversification, and the regression analysis techniques (OLS and QR).

Aggregation of Firms into Groups

The sample contains firms that belong to 70 different real sector industries. To simplify the analysis an aggregation into eight industrial groups was done. The firms were discriminated by their OKVED industry code keeping close to NACE high-level SNA aggregation principles. The details are provided in Table 1 below.

Table 1

Aggregated Industrial Groups

Group Number and Encoding	Group Name	OKVED Codes
№1 (A)	Agriculture, Fishing, and Forestry	01+02+03
№2 (B)	Mining	05+06+07+08+09
№3 (C)	Manufacturing	10+11+12+13+14+15+16+17+18+19+20+ +21+22+23+24+25+26+27+28+29+30+31+ +32+33+35+36+37+38+39
№4 (D)	Wholesale, Retail and Food Services	45+46+47+55+56
№5 (E)	Construction	41+42+43+68
№6 (F)	Transport	49+50+51+52
№7 (G)	IT, Broadcasting, and Telecommunication	53+58+59+60+61+62+63
№8 (H)	Services and Other Minor Industries	69+71+73+77+78+79+81+82+95+96

Note. Mining relates to any natural resource extraction activity, not only coal mining.

Descriptive Statistics

There are three key variables used in the paper: *employees* that was used directly in the model and as an input factor to calculate EG and HHI, and performance measures: *sales margin* and *profit margin*. Discussion on measuring performance is provided in the *Modeling Techniques* subsection. The present section contains a brief overview of the descriptive statistics of those key variables (Table 2).

Table 2

Descriptive Statistics for the Industry Groups

		Group A	Group B	Group C	Group D	Group E	Group F	Group G	Group H
Sales Margin	*Mean*	19.83	24.31	17.59	15.72	21.31	16.92	24.96	23.55
	Median	14.47	17.87	10.93	10.00	11.16	7.69	15.10	14.06
	Std.Dev.	18.59	21.83	18.94	17.25	24.10	22.24	25.19	24.26
	Min	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	Max	99.99	99.93	100.00	100.00	100.00	100.00	100.00	100.00
Net Profit Margin	*Mean*	14.60	9.24	7.28	5.36	10.93	7.69	12.92	13.20
	Median	9.79	4.16	3.10	2.10	3.51	2.21	5.97	5.82
	Std.Dev.	18.73	16.91	13.65	11.57	19.57	15.07	19.28	20.00
	Min	–98.53	–95.87	–99.67	–99.90	–99.94	–96.40	–99.56	–99.76
	Max	99.68	91.59	99.98	99.89	99.98	99.39	99.88	99.96
Employees	*Mean*	37.91	163.60	40.84	9.32	12.17	23.12	14.93	9.82
	Median	9.00	14.00	8.00	4.00	5.00	5.00	5.00	4.00
	Std.Dev.	106.06	1,034.72	207.61	29.51	40.53	182.71	70.38	38.40
	Min	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
	Max	3,184.00	40,250.00	19,940.00	3,977.00	4,333.00	13,010.00	7,648.00	4,360.00
Observations		20,116	2,509	89,952	276,171	128,322	34,047	23,773	72,807

Note. Group A — Agriculture, Fishing, and Forestry, Group B — Mining, Group C — Manufacturing, Group D — Wholesale, Retail and Food Services, Group E — Construction, Group F — Transport, Group G — IT, Broadcasting, and Telecommunication, Group H — Services and Other Minor Industries.

The standard deviation is comparable to the mean and median values for performance measures and is drastically higher for *employees*. The reason for such deviation in the number of employees is straightforward: in each group, companies of different scales were included. Such fluctuations do not influence the precision of indices (industry localization and regional economic diversification) and can be considered in the regression analysis via the introduction of dummy variables. Descriptive statistics also show that the industries differ significantly, which supports the idea of industry division and separate analysis.

All the key variables do not seem to be normally distributed according to the data in Table 2. However, their logarithms do, which were indeed the model inputs.

The dependent variables are highly dispersed, and this might be a problem as the effects driven by localization and diversification might be different for high and low margin firms. This phenomenon is studied here by applying QR analysis (which allows one to consider each quantile of the dependent variable separately).

Industry Localization and Regional Economic Diversification Measures

To measure industry localization, EG [Dumais et al., 1997] was applied. This index may be considered as a robust measure of industry localization and works well for different industries [Cassey, Smith, 2014]. Regional diversity was measured with the HHI, which is a simple and widespread measure of diversification. Calculation issues are presented in Table 3.

Table 3

Indices of Industry Localization and Regional Economic Diversification

Index	Calculation Issues	Description
EG for Measuring Industry Localization	$\gamma_j = \frac{\Sigma_i (s_{ij} - x_i)^2 - (1 - \Sigma_i x_i^2)\Sigma_j z_j^2}{(1 - \Sigma_i x_i^2)(1 - \Sigma_j z_j^2)}$ EG is unbounded either above or below. Lower values correspond to the situation of random spatial distribution of companies; higher values stand for localized industries in certain region(s)	$j = 1…70$ – industries $i = 1…82$ – Russian regions s_{ij} – share of industry j in region i in terms of employment x_i – share of region i in the whole country's employment z_j – share of industry j in the whole country's employment
HHI for Measuring Regional Economic Diversification	$HHI_{s_i} = \sum_j (s_{ij})^2$ $HHI_{s_i} \in [\frac{1}{j}; 1]$ Lower values – regional economy is diversified Higher values – regional economy is specialized	$j = 1…70$ – industries $J = 70$ – total number of industries (disaggregated) $i = 1…82$ – Russian regions s_{ij} – share of industry j in region i in terms of employment

Other diversification measures such as Theil and Gini indices or Vorobyev modification of the HHI [Vorobyev et al., 2010] are not being used in the present paper. They range the regions by the degree of economy diversity almost the same way (rank correlation between sorted lists of regions is higher than 0.9), and therefore their simultaneous inclusion in the regression is fruitless. Moreover, the HHI is the most widespread and simple measure of regional diversification, and it turns out to be robust when working with general population proxies. Localization can also be measured in absolute terms through the geocoding procedure [Aleksandrova et al., 2020], although such an operation requires considerable computational power and the results still need to be classified in order to be included in the regression equation. Thus, preference was given to the relative localization indicator—the EG index.

There is no direct linkage between EG and HHI. The fact that industry *j* is localized in some region *i* does not necessarily make this region specialized. The economy of region *i* may be diversified, for example, if industry *j* is small or the region's economy is developed enough. Therefore, it is important to look at both indices and study the effects of industry localization and regional economic diversity separately.

Both indices indirectly measure agglomeration effects and they are fading for the firms located far enough from cities [Rosenthal, Strange, 2004]. However, in this paper the effects are studied at the regional level. This might be approved as [Ciccone, Hall, 1996; Davidson, Mariev, 2015; Henderson, 2003] showed that the agglomeration effects could still be identified and measured, and influence the enterprises at the regional and small country levels.

The estimation of the indices is endogenous and based on the available data about the employees, OKVED codes and regional codes.

Modeling Techniques

The research question is whether and how the degree of industry localization in the region and regional economic diversity influence firms' performance indicators. High dispersion of the dependent variables and consequent interest, whether localization and diversification effects are the same for different implied performance firms, makes justified the estimation of the two different models: OLS and QR for each industry group separately. This section provides information about the specification of both models.

In the present paper the firms' performance was measured with sales margin (equally *"SM"*) and net profit margin (or *"NPM"*) indicators. This choice contrasts with the approach applied in the closest (in terms of methodology) paper [Davidson, Mariev, 2015], where revenue was the dependent variable. Revenue absorbs greater among-firms variance; however, it is an absolute measure and fails to distinguish between an inefficient giant enterprise and a small competitive company. Markup was not chosen either as it requires much approximation due to the lack of available data (input prices or elasticities) [Lu et al., 2012]. Efficiency indicators are likely to be good proxies to the firm's performance, and some detailed discussion about them follows below. Another approach that has worked well on Russian data is to use the revenue per unit of labor ratio [Lavrinenko et al., 2019]. However, the sample consists of a wide range of companies with a heterogeneous structure of costs, and precise clarification of labor contribution to the cost function is not possible.

$$SM = \frac{Revenue - Costs\ of\ sales}{Revenue}, \quad (1)$$

$$NPM = \frac{Net\ profit}{Revenue}. \quad (2)$$

NPM might look more interesting as it directly reflects final stakeholders' returns, while SM only measures the interim results. However, NPM has several pitfalls and distortions. It might seriously vary from year to year depending highly on the firm's accounting policy. Net profit sometimes includes deferred profit, tax deductions (VAT deductions, for instance) and a lot of other information that can change the intrinsic performance during the current year away from the P&L accounting sheet.

From this perspective, SM looks more representative as it only includes information about revenues and costs from the main declared activity in the current financial year and is not affected by other financial report lines. Nevertheless, it is effective to look at both performance indicators to compare and control the results.

Call the dependent variable Y regardless of the type (SM or NPM). The assumption of a normally distributed error term is excessive in this case as the high number of observations guarantees that coefficients would converge into normally distributed estimates $\hat{\beta} \longrightarrow N$. The OLS model in its general form can be written as:

$$Y = \beta_0 + X_1\beta_1 + X_2\beta_2 + X_3\beta_3 + \varepsilon, \quad (3)$$

where

β_k – regression coefficients ($k = 1…3$, β_0 – intercept);
X_1 – firm specific regressors;
X_2 – region specific regressors;
X_3 – matrix of explanatory variables related to industry localization and regional economic diversification;
ε – vector of random errors.

QR in its general form has the same structure:

$$Q_\tau(Y) = \beta_0(\tau) + X_1\beta_1(\tau) + X_2\beta_2(\tau) + X_3\beta_3(\tau) + \varepsilon, \tag{4}$$

where

τ – corresponding quantile;
$\beta_k(\tau)$ – regression coefficients for quantile τ ($k = 1 \dots 3$, β_0 – intercept);
X_1 – firm specific regressors;
X_2 – region specific regressors;
X_3 – matrix of explanatory variables related to industry localization and regional economic diversification;
ε – vector of random errors.

The major difference is in the estimation procedure. Coefficients in the QR model are estimated with numerical methods. In the present paper, the Frisch-Newton interior point method of estimation was applied as the sample is too large for other built-in algorithms. In Table 4 all the regressors included in both models are described.

Table 4

Variables That Were Included in the OLS and QR Models

Variable	Variable Description	Reasons for Inclusion	Expected Effects
	Dependent Variable		
NPM	Net profit margin	Related to the main hypothesis	NA
SM	Sales margin	Related to the main hypothesis	NA
	Enterprise Specific Variables (X_1 Regressors)		
logNetAss	Log of the firm's net assets. Net assets equals capital + + future receivables – – financial assets	Industrial scale measure (the higher it is the greater is the value of assets owned by the firm)	+
logEmployees	Log of enterprise annual weighted average employment	Industrial scale measure (proxy to labor input costs)	–
logDEBTTOGRPof	Net debt/gross profit multiplier. Net debt is the sum of long-term and short-term debts subtracting cash and its equivalents. Gross profit is calculated as the difference between revenue and cost of sales	Measure of debt burden (the higher it is, the greater is the interest paid and the lower would NPM be)	–
Firm scale Dummies	Dummy indicating enterprise scale (big, medium, small and micro companies)	Measure of cash flow	Two-side effects
	Region Specific Variables (X_2 Regressors)		
logAverMonthWage	Log of the average accrued wage in a region	Measure of labor productivity; production factor (labor) price	Two-side effects

Continuation of Table 4

Variable	Variable Description	Reasons for Inclusion	Expected Effects
logRND	Log of R&D investments	Proxy to internal regional investments; measure of labor quality and labor factor productivity	+
Unempl	Regional unemployment rate	Proxy measure of regional economic activity	Inverse U-shape effect
Unemplsq	Squared regional unemployment rate	Proxy measure of regional economic activity	Inverse U-shape effect
UnprofitWeight	Share of companies with losses in a region	Proxy measure of regional economic activity; proxy to competition intensity	Two-side effects
DiffInvest	Percentage increase of investments in a region compared to the previous year (2016)	Proxy to incoming regional investments	+
FirmBirthRate	Ratio of the difference between created and closed companies in 2017 to the total number of firms in a region at the beginning of 2017	Proxy measure of regional economic activity	+
CriminalRate	Number of registered crimes per 100,000 in a region	Proxy measure of institutional climate quality and "doing business" environment	–
Industry Localization and Regional Economy Diversification (X_3 Regressors)			
Ellison-Glaeser	Ellison-Glaeser industry localization index	Related to the main hypothesis	+/– depending on industry
HHIs	Regional diversification measured by the simple Herfindahl-Hirschman index	Related to the main hypothesis	+/– depending on industry
logcore	Log of the ratio of the firm's industry employment to total regional employment (industry share in a region)	Related to the main hypothesis	+/– depending on industry

Possible omitted variable bias was controlled at least at the stage of defining variables' powers up to the 3rd in the model. To guarantee the inclusion of the necessary powers, the Ramsay test was applied.

Finally, this paper does not provide justification for whether localization and diversification have a causal relation to firm performance, or the obtained effects are just random correlations. However, since there is a strong theoretical basis that has been developing during the past century and includes detailed description of the channels of how the studied variables can affect each other, there are reasons to treat the results as causal.

5. Results

The complex nature of the research question and stated hypotheses requires the consideration of two performance indicators (SM and NPM), eight industrial groups, and two methodologies (OLS and QR). Thus, the output is an ample set that consists of 16 OLS models (2*8) and 144 QR models (2*8*9, where 9 is the number of considered quantiles). For convenience reasons the discussion of the estimation results was split into two separate subsections.

OLS

The estimation results for the model that measures performance via sales margin are provided in Table 5, whereas Table 6 represents the output of the model with net profit margin dependent variables.

No influence of either the degree of industry localization or the degree of regional economy diversification was found for mining industry enterprises. This is not a surprising finding as Russian mining companies' distribution chains usually go far beyond the borders of their home region—exporting natural resources to the other Russian regions or even abroad. For this reason, *Mining* companies initially seemed unlikely to depend on the internal markets, and therefore on the degree of regional economy diversification. Localization also does not affect the performance, since *Mining* industry companies are localized around natural resource fields in limited numbers (as a mining license is required), so their location decisions are naturally rather than economically driven.

Regional diversification generates negative effects for companies of all industrial groups except *Agriculture, Fishing, and Forestry*, where some positive trends for the net profit margin were found. This might be explained by the following logic: *Sales, Manufacturing* and *Transport* are the largest in terms of employment in the Russian economy's industries, and thus an increase in the degree of diversification might often mean a higher weight of other companies in regional employment, which may lead to increasing competition between them. That channel might finally result in a market markup reduction the same way as was found for Chinese companies [Lu et al., 2012]. However, the results seem to contradict those obtained by Davidson and Mariev. There are two possible explanations. Technically the contradiction could be connected to the differences in the approaches to measurement of companies' efficiency. Moreover, a 9-year gap is large enough to the economy to contain some structural shifts. To verify this, an additional research conducted on a larger 2002–2017 panel is necessary.

The most powerful influence of the degree of regional economy diversification on the sales margin was found in *Agriculture, Fishing, and*

Forestry firms. The same indicator influences the net profit margin inversely for the firms within these industrial groups. This could be related to various non-market operations or internal operations that were not included in the analysis.

Localization effects are heterogeneous for companies from different industrial groups. Localization enhances the performance of *Agriculture, Fishing, and Forestry; Manufacturing; Trade and Food Services; Transport;* and *IT, Broadcasting, and Telecommunication* companies, and reduces the margin of *Construction* and *Services* industry companies. The most significant is the influence of localization on the *Building and Construction* companies: a 0.1 EG change reflects a more than 7-fold percentage decrease in the sales margin. In the same direction the degree of localization changes the net profit margin of *Construction* industry companies: a 0.1 index change leads to a 1.3-factor performance drop. The positive effects verified here are in line with the existing results [Davidson, Mariev, 2015; Holmes, 1999; Martin et al., 2011].

The divergence from the earlier findings is likely to be connected to separate consideration of different industries here. One suggested explanation for such an influence of localization in the *Construction* industry may lie in the problem of ground shortage in certain local zones. *Construction* industry firms depend highly on the quality of placement. Localization may mean an excess of construction industry companies in the region and thus increasing competition for the ground and, at the same time, a fall in the number of consumers per firm. Growing costs and decreasing demand result in diminishing profits and performance indicators.

On the contrary, *Transportation* and *Agriculture, Fishing, and Forestry* groups of industries have the largest gains from industry localization both for sales margin and net profit margin.

One possible explanation, applicable for the *Agriculture, Fishing, and Forestry* industrial group, might be related to the quality of the areas shown. In this case the companies are naturally localized in places where the best, rich and stable yield is possible. Despite the high ground rent costs, localization is justified due to the increase in production and in the quality of the yield. It seems that the positive effects from the increase in productivity outperform the negative ones brought about by the increase in rental costs.

For the *Transportation* industry, an increase in the degree of localization might be associated with the presence of big cities within a region where the importance of the industry increases. Ticket prices are also comparably higher, which opens an opportunity for extracting higher market markups.

The bottom line is as follows: simultaneous existence of localization and diversification effects was found for all the industries except *Min-*

Table 5

Results of OLS Estimation of Sales Margin on Internal and Regional Factors for Firms Divided by Industry Groups

	Agriculture, Fishing, and Forestry	Mining	Manu-facturing	Wholesale, Retail and Food Services	Construction	Transport	IT, Broadcasting, and Telecom-munication	Services and Minor Industries
(Intercept)	2.5283*** (0.3975)	2.5423** (0.9334)	2.2316*** (0.1390)	2.1781*** (0.0793)	1.4525*** (0.1250)	2.6202*** (0.2478)	3.0690*** (0.2768)	2.8750*** (0.1618)
logNetAss3	0.1705*** (0.0054)	0.2347*** (0.0141)	0.2444*** (0.0023)	0.2532*** (0.0013)	0.2968*** (0.0018)	0.3137*** (0.0037)	0.3018*** (0.0051)	0.3213*** (0.0026)
logEmployees	−0.1987*** (0.0101)	−0.2040*** (0.0251)	−0.2434*** (0.0044)	−0.1329*** (0.0029)	−0.2844*** (0.0040)	−0.2774*** (0.0078)	−0.2866*** (0.0092)	−0.3354*** (0.0055)
logDEBTTOGRPof	−0.2565*** (0.0060)	−0.2723*** (0.0196)	−0.2768*** (0.0033)	−0.2660*** (0.0021)	−0.1914*** (0.0032)	−0.2474*** (0.0060)	−0.3156*** (0.0113)	−0.2581*** (0.0056)
BigFirm	0.1000* (0.0498)	−0.0145 (0.0809)	0.0583** (0.0208)	0.1051*** (0.0156)	−0.0478 (0.0376)	0.0359 (0.0617)	−0.0926 (0.0748)	0.0530 (0.0694)
SmallFirm	−0.2716*** (0.0359)	−0.0012 (0.0740)	−0.1575*** (0.0160)	−0.1238*** (0.0105)	0.0339 (0.0246)	−0.1163** (0.0437)	−0.1723** (0.0529)	−0.1050* (0.0468)
MicroFirm	−0.4266*** (0.0432)	−0.0952 (0.1026)	−0.3082*** (0.0192)	−0.1449*** (0.0119)	0.1622*** (0.0259)	−0.1630*** (0.0462)	−0.3236*** (0.0545)	−0.1333** (0.0464)
logAverMonthWage	−0.0188 (0.0380)	−0.0452 (0.0813)	0.1176*** (0.0128)	0.1104*** (0.0068)	0.0700*** (0.0108)	0.1097*** (0.0214)	0.1039*** (0.0240)	0.0651*** (0.0138)
logRND	0.0143*** (0.0037)	−0.0030 (0.0102)	−0.0007 (0.0016)	−0.0049*** (0.0010)	0.0058*** (0.0016)	−0.0085** (0.0031)	−0.0058 (0.0040)	−0.0013 (0.0021)
Unempl	−0.0167 (0.0116)	0.0384 (0.0323)	−0.0050 (0.0056)	−0.0021 (0.0034)	−0.0147** (0.0047)	0.0039 (0.0123)	0.0211' (0.0123)	0.0061 (0.0072)
Unemplsq	−0.0010 (0.0006)	−0.0024 (0.0017)	−0.0002 (0.0003)	−0.0002 (0.0002)	0.0002 (0.0002)	−0.0009 (0.0009)	−0.0015* (0.0007)	−0.0006 (0.0004)
UnprofitWeight	0.0063*** (0.0015)	0.0130** (0.0045)	0.0003 (0.0008)	0.0020*** (0.0005)	0.0044*** (0.0008)	0.0043** (0.0016)	−0.0032 (0.0020)	0.0010 (0.0011)

Continuation of Table 5

	Agriculture, Fishing, and Forestry	Mining	Manu-facturing	Wholesale, Retail and Food Services	Construction	Transport	IT, Broadcasting, and Telecom-munication	Services and Minor Industries
DiffInvest	–0.0004 (0.0004)	0.0013 (0.0011)	–0.0005* (0.0002)	–0.0010*** (0.0001)	–0.0002 (0.0002)	0.0003 (0.0004)	–0.0007 (0.0005)	–0.0014*** (0.0003)
FirmBirthRate	–0.0012** (0.0005)	0.0008 (0.0016)	–0.0008** (0.0002)	0.0000 (0.0001)	–0.0007** (0.0002)	–0.0008' (0.0005)	–0.0005 (0.0006)	–0.0003 (0.0003)
CriminalRate	0.0001** (0.0000)	0.0000 (0.0001)	0.0000 (0.0000)	–0.0000*** (0.0000)	–0.0000 (0.0000)	–0.0001*** (0.0000)	–0.0001*** (0.0000)	–0.0000 (0.0000)
EG	***5.7067****** ***(0.3923)***	***0.1803*** ***(0.1703)***	***1.8520****** ***(0.4003)***	***5.4161****** ***(0.7154)***	***–70.7048****** ***(2.3383)***	***13.8603****** ***(2.0790)***	***0.1266****** ***(0.0379)***	***–0.9954****** ***(0.1049)***
HHIs	***0.6518**** ***(0.2579)***	***0.6276*** ***(0.5211)***	***0.0032*** ***(0.0750)***	***0.1295***** ***(0.0467)***	***0.2302****** ***(0.0697)***	***0.3675**** ***(0.1464)***	***–0.0212*** ***(0.1803)***	***0.2221**** ***(0.0977)***
logcore	0.0602*** (0.0078)	–0.0335* (0.0136)	0.0014 (0.0031)	–0.0674*** (0.0033)	–0.1268*** (0.0091)	0.0513*** (0.0098)	0.0704*** (0.0069)	0.0373*** (0.0044)
R^2	0.18	0.25	0.22	0.27	0.28	0.31	0.20	0.25
Adj. R^2	0.18	0.25	0.22	0.27	0.28	0.31	0.20	0.25
Num. obs.	20,116	2,509	89,952	276,171	128,322	34,047	23,773	72,807

Note: Significance level: ‘— $p < 0.1$, *— $p < 0.05$, **— $p < 0.01$, ***— $p < 0.001$.

Table 6

Results of OLS Estimation of Net Profit Margin on Internal and Regional Factors for Firms Divided by Industry Groups

	Agriculture, Fishing, and Forestry	Mining	Manu-facturing	Wholesale, Retail and Food Services	Construction	Transport	IT, Broadcasting, and Telecom-munication	Services and Minor Industries
(Intercept)	4.4594*** (0.0712)	4.7088*** (0.1467)	4.6466*** (0.0191)	4.5972*** (0.0098)	4.6910*** (0.0206)	4.7291*** (0.0290)	4.7433*** (0.0466)	4.7523*** (0.0261)
logNetAss3	0.0266*** (0.0010)	0.0204*** (0.0028)	0.0214*** (0.0004)	0.0170*** (0.0002)	0.0276*** (0.0003)	0.0277*** (0.0005)	0.0359*** (0.0009)	0.0381*** (0.0005)
logEmployees	–0.0332*** (0.0019)	–0.0333*** (0.0044)	–0.0277*** (0.0007)	–0.0212*** (0.0004)	–0.0390*** (0.0007)	–0.0341*** (0.0011)	–0.0488*** (0.0017)	–0.0550*** (0.0009)
logDEBTTOGRPof	–0.0270*** (0.0013)	–0.0246*** (0.0035)	–0.0191*** (0.0006)	–0.0122*** (0.0004)	–0.0243*** (0.0007)	–0.0165*** (0.0010)	–0.0264*** (0.0019)	–0.0291*** (0.0012)
BigFirm	0.0086 (0.0092)	0.0251* (0.0117)	0.0025 (0.0021)	–0.0024* (0.0010)	–0.0073 (0.0052)	0.0154* (0.0073)	–0.0121 (0.0146)	0.0001 (0.0089)
SmallFirm	–0.0018 (0.0064)	–0.0280** (0.0105)	–0.0049** (0.0016)	0.0027*** (0.0007)	–0.0033 (0.0026)	–0.0129** (0.0040)	–0.0044 (0.0112)	0.0093 (0.0060)
MicroFirm	–0.0111 (0.0075)	–0.0234 (0.0170)	0.0064** (0.0025)	0.0097*** (0.0012)	0.0162*** (0.0031)	–0.0056 (0.0047)	0.0152 (0.0116)	0.0195** (0.0062)
logAverMonthWage	0.0192** (0.0068)	–0.0064 (0.0132)	0.0041* (0.0017)	0.0003 (0.0008)	–0.0056** (0.0018)	0.0023 (0.0024)	0.0045 (0.0041)	–0.0001 (0.0022)
logRND	0.0026*** (0.0006)	–0.0004 (0.0019)	–0.0007** (0.0002)	0.0005*** (0.0001)	0.0003 (0.0003)	–0.0013*** (0.0004)	–0.0003 (0.0006)	–0.0007' (0.0004)
Unempl	0.0017 (0.0018)	0.0053 (0.0046)	0.0003 (0.0007)	–0.0015** (0.0005)	–0.0005 (0.0008)	0.0037** (0.0014)	0.0093*** (0.0020)	0.0006 (0.0012)
Unemplsq	–0.0002* (0.0001)	–0.0002 (0.0002)	0.0000 (0.0000)	0.0002*** (0.0000)	0.0001 (0.0000)	–0.0002* (0.0001)	–0.0004*** (0.0001)	–0.0000 (0.0001)
UnprofitWeight	0.0005' (0.0003)	0.0011 (0.0008)	0.0004*** (0.0001)	0.0006*** (0.0001)	0.0005*** (0.0001)	0.0007** (0.0002)	–0.0004 (0.0003)	0.0003 (0.0002)

Continuation of Table 6

	Agriculture, Fishing, and Forestry	**Mining**	**Manu-facturing**	**Wholesale, Retail and Food Services**	**Construction**	**Transport**	**IT, Broadcasting, and Telecom-munication**	**Services and Minor Industries**
DiffInvest	–0.0001 (0.0001)	–0.0001 (0.0002)	–0.0000 (0.0000)	–0.0000* (0.0000)	0.0000 (0.0000)	–0.0000 (0.0000)	–0.0001* (0.0001)	–0.0000 (0.0000)
FirmBirthRate	–0.0001 (0.0001)	0.0000 (0.0002)	–0.0000 (0.0000)	0.0001*** (0.0000)	–0.0000 (0.0000)	0.0001 (0.0001)	0.0000 (0.0001)	–0.0000 (0.0001)
CriminalRate	0.0000** (0.0000)	0.0000 (0.0000)	0.0000*** (0.0000)	0.0000*** (0.0000)	0.0000 (0.0000)	–0.0000 (0.0000)	–0.0000** (0.0000)	–0.0000 (0.0000)
EG	0.3304** (0.1010)	0.0273 (0.0318)	–0.0736 (0.0539)	0.5796*** (0.1037)	–13.4015*** (0.4195)	1.6030*** (0.3111)	0.0428*** (0.0053)	–0.1151*** (0.0165)
HHIs	–0.1425** (0.0445)	–0.0784 (0.1230)	0.0233* (0.0099)	0.0169** (0.0057)	0.0335** (0.0117)	0.0154 (0.0167)	***0.0562* (0.0282)***	***0.0257 (0.0168)***
logcore	0.0083*** (0.0015)	–0.0004 (0.0022)	–0.0004 (0.0004)	–0.0121*** (0.0005)	–0.0139*** (0.0015)	0.0073*** (0.0013)	0.0125*** (0.0011)	0.0042*** (0.0008)
R^2	0.09	0.09	0.09	0.08	0.13	0.16	0.12	0.14
Adj. R^2	0.09	0.08	0.09	0.08	0.13	0.16	0.12	0.14
Num. obs.	20,116	2,509	89,952	276,171	128,322	34,047	23,773	72,807

Note: Significance level: ‘— $p < 0.1$, *— $p < 0.05$, **— $p < 0.01$, ***— $p < 0.001$.

ing. *Agriculture, Fishing, and Forestry* is the only industry where both effects turned out to be positive, whereas *Construction* and *Other Services* are the industries that suffer both from an increase in EG and from a decrease in HHI. The rest of the companies gain from localization, and diversification decreases their margins. Thus, the first hypothesis *(a.)* cannot be rejected as agglomeration effects were established and these effects vary depending on the industry.

Finally, a few comments concerning the control variables. Higher net assets mean better leverage possibilities for the firm as the amount of company-owned assets increases. Being debt-free, these assets generate a higher margin. The opposite case is an increase in debt burden, which leads to higher interest paid and lower net profit margins. A large number of employees often leads to inefficiencies in business processes, which reflects in negative effects on either SM or NPM for any company. Larger company scale means better market positions and opportunities due to the increasing monopolistic power associated with greater markup. Better wages are associated with greater productivity of labor that leads to the generation of a higher margin.

Quantile Regression

The companies differ significantly, not only in terms of scale or industry, but also in terms of their business profile within their own categories. Performance ranges significantly from –86% to 100%, and it is natural to check whether the localization and diversification effects stay the same for companies from different quantiles. The modeling logic stays the same: for each industry group, its own regression is estimated. The observed percentiles range from 10% to 90% with a 10% step.

Estimation results for the main variables (degree of localization and regional economy diversification) are presented graphically in Figures 1a–1h, 2a–2h, 3a—3h, and 4a–4h. Complete quantile regression outputs are openly available[7].

For most industries, the results stayed in line with the OLS. For example, it is notable that, for the *Mining* industry, QR coefficients of any quantile stayed inside the OLS confidence bands. However, there are exceptions—industries where quantile analysis allows one to extend the knowledge concerning the linkage between the degree of industry localization and performance indicators (Group C – *Manufacturing*, Group D – *Wholesale, Retail and Food Services*, Group E – *Construction*, and Group H – *Services and Other Minor Industries*).

Keeping in mind that higher quantiles represent more efficient businesses, one may notice that industry localization greatly enhances the

[7] The regression output tables are available in open access at the link below. For convenience purposes, they were replaced by pictures in the paper text. https://www.dropbox.com/sh/jua0fwseo9r7688/AADB-GLg8M24S8bkxn9NLnK_Da?dl=0 (Tables 7–10).

performance of low margin manufacturing firms, and the effect fades for the industry leaders (becoming negative for the best performing companies). The results are in line with the widespread idea [Audretsch, Feldman, 1996; Jacobs, 1961; Shaver, Flyer, 2000] that smaller and generally poorer firms benefit more from agglomeration and bigger ones lose. The more successful firms are general donors of ideas and technologies through spillover channels, while smaller or less successful firms simply deploy ready ideas without spending extra sums on trials and errors.

Regardless of the quantile, industry localization generates negative effects for the *Construction* industry, and the explanation for this fact is the same as was provided for the OLS model. Moreover, the negative effect increases for higher margin firms. One suggested explanation might be connected to the fact that highest margins are generated when dealing with premium market segments, which requires even more expensive building areas. High localization circumstances drive the prices up even more, and the margin drops even more dramatically. This situation is especially common for big cities like Moscow or Saint Petersburg.

The power of the benefits from industry localization increases along with the performance for companies that work in the *Wholesale, Retail, and Food Services* industry. A possible explanation suggests that there are no causal effects here. The reason, again, is market segmentation, but it is passive in this instance. In contrast to the construction industry, where building area costs might be crucial, for trade (*Wholesale, Retail and Food Services*) industries, major costs are associated with the goods produced, and localization just means the availability of the supply side. Good transportation nodes for a logistic wholesale company or restaurant placement generate long-term benefits in terms of increasing margin, while for the construction industry placement costs are immediate sunk costs.

The effects of regional economy diversification turned out to be less influential and generally repeat the results obtained with the OLS. The linkage between the regional diversification degree and firms' performance is weaker than between localization and performance. The channels that connect regional diversification and margin-adding processes turned out to be weak at the regional level. However, the analysis contains opportunities for future research with the presence of big cities in a region considered, where stronger effects might be found.

The second *(b.)* hypothesis stated that localization and diversification effects depend not only on the industry, but also on the firm's implied business effectiveness. It cannot be rejected due to the findings described above. Moreover, as has been suggested, these effects appeared to be non-monotone, and sometimes U-shaped or even W-shaped dependencies were discovered.

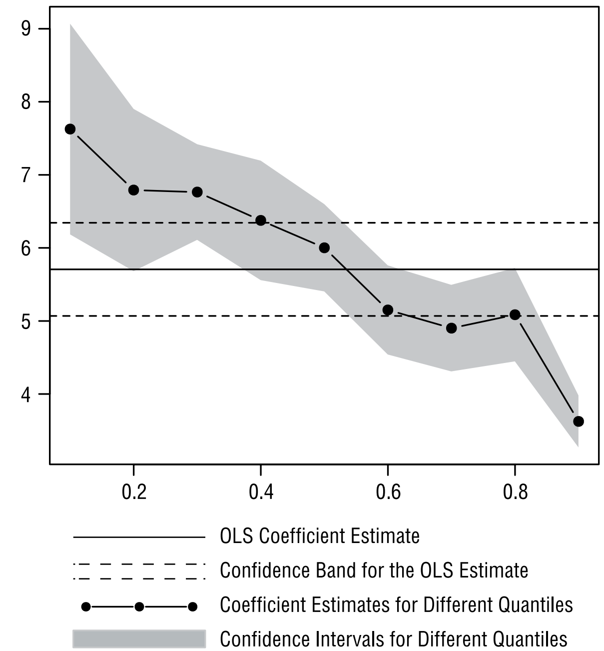

Fig. 1a. **Beta Coefficients of EG Regressor** (Y-Axis) **for Different Quantiles** (X-Axis) **of the Agriculture, Fishing, and Forestry Industry in the Model with Sales Margin Dependent Variable**

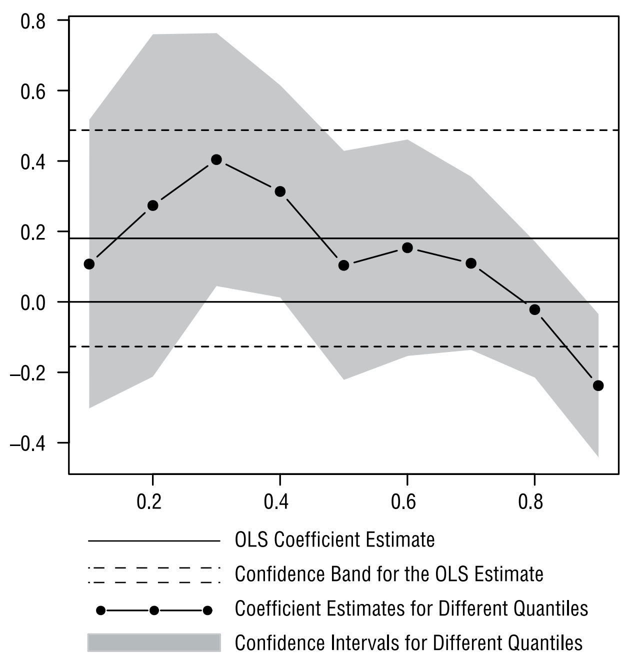

Fig. 1b. **Beta Coefficients of EG Regressor** (Y-Axis) **for Different Quantiles** (X-Axis) **of the Mining Industry in the Model with Sales Margin Dependent Variable**

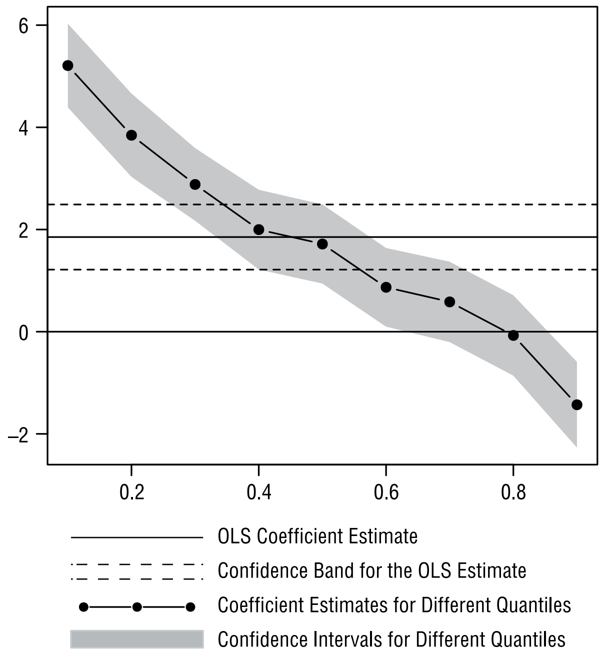

Fig. 1c. **Beta Coefficients of EG Regressor** (Y-Axis) **for Different Quantiles** (X-Axis) **of the Manufacturing Industry in the Model with Sales Margin Dependent Variable**

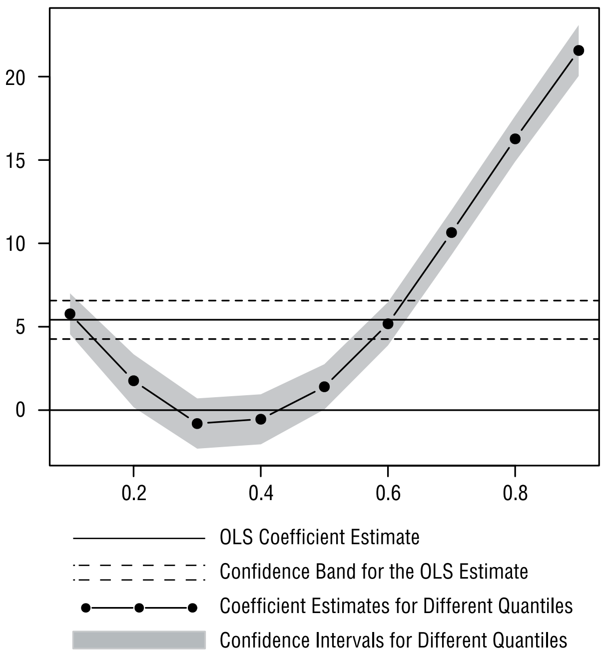

Fig. 1d. **Beta Coefficients of EG Regressor** (Y-Axis) **for Different Quantiles** (X-Axis) **of the Wholesale, Retail and Food Services Industry in the Model with Sales Margin Dependent Variable**

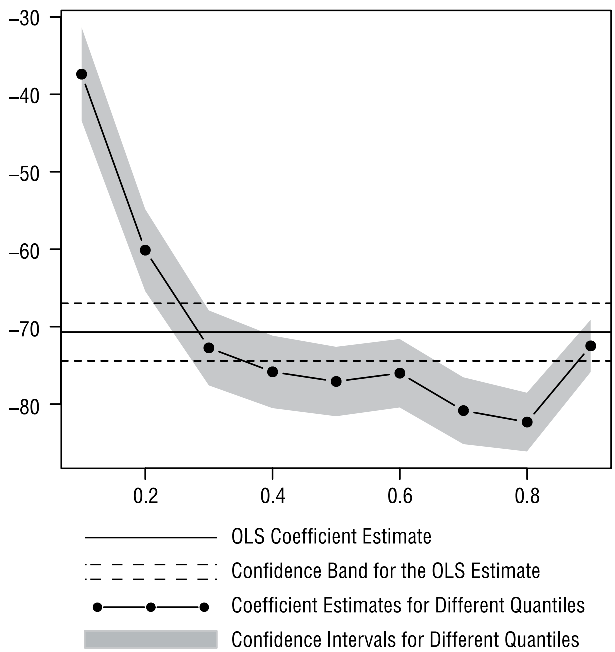

Fig. 1e. **Beta Coefficients of EG Regressor** (Y-Axis) **for Different Quantiles** (X-Axis) **of the Construction Industry in the Model with Sales Margin Dependent Variable**

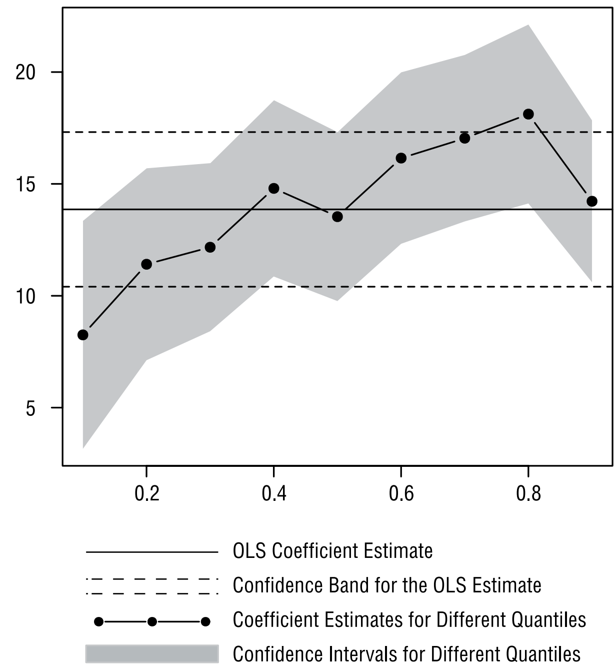

Fig. 1f. **Beta Coefficients of EG Regressor** (Y-Axis) **for Different Quantiles** (X-Axis) **of the Transport Industry in the Model with Sales Margin Dependent Variable**

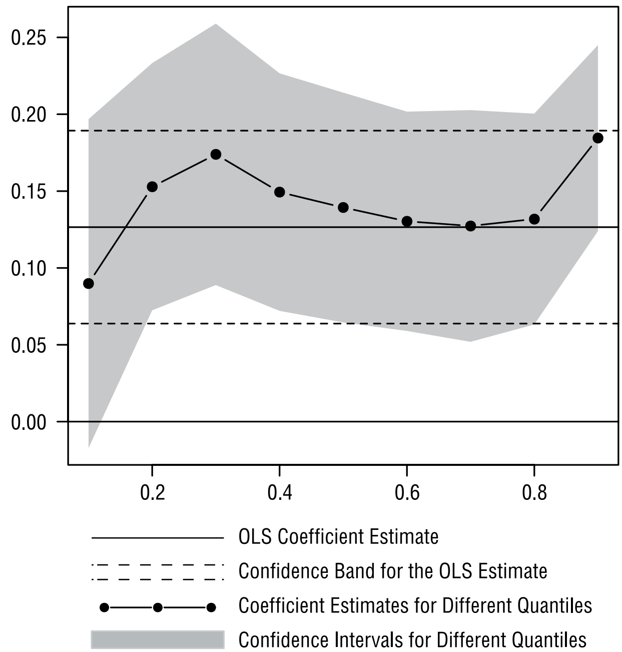

Fig. 1g. **Beta Coefficients of EG Regressor** (Y-Axis) **for Different Quantiles** (X-Axis) **of the IT, Broadcasting, and Telecommunication Industry in the Model with Sales Margin Dependent Variable**

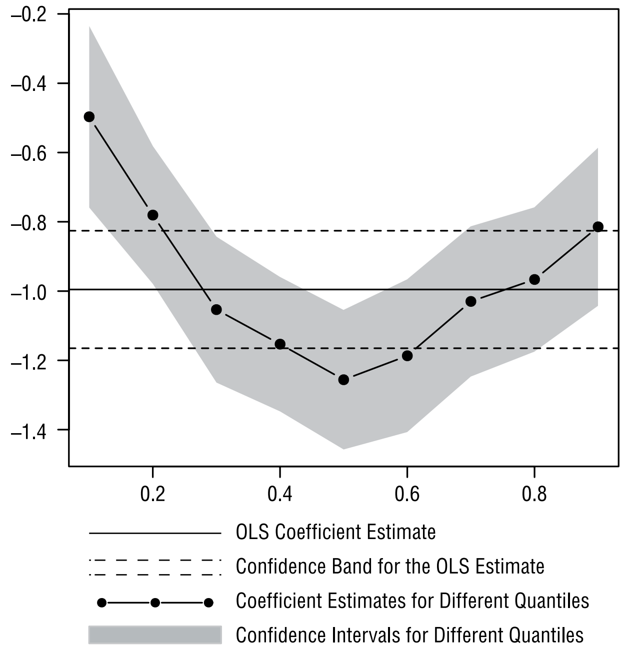

Fig. 1h. **Beta Coefficients of EG Regressor** (Y-Axis) **for Different Quantiles** (X-Axis) **of the Services and Minor Industries Industry in the Model with Sales Margin Dependent Variable**

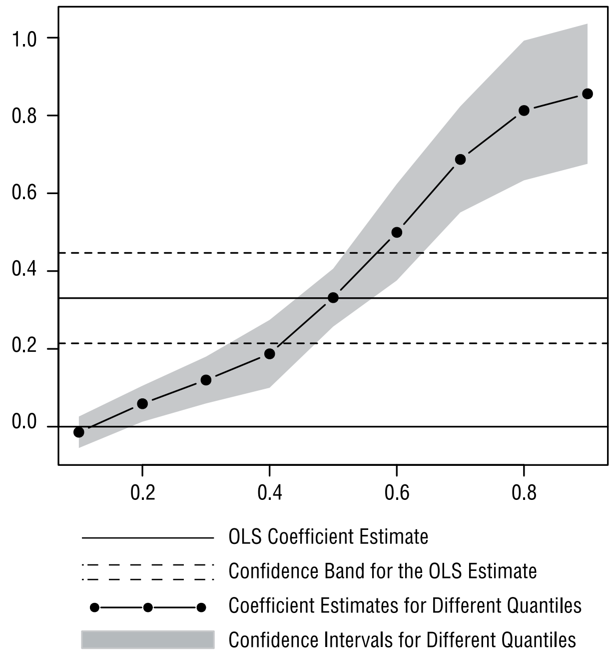

Fig. 2a. **Beta Coefficients of EG Regressor** (Y-Axis) **for Different Quantiles** (X-Axis) **of the Agriculture, Fishing, and Forestry Industry in the Model with Net Profit Margin Dependent Variable**

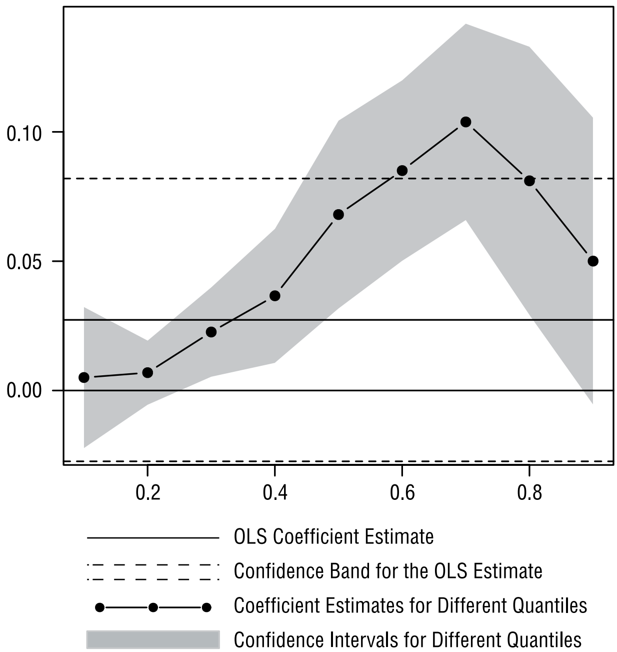

Fig. 2b. **Beta Coefficients of EG Regressor** (Y-Axis) **for Different Quantiles** (X-Axis) **of the Mining Industry in the Model with Net Profit Margin Dependent Variable**

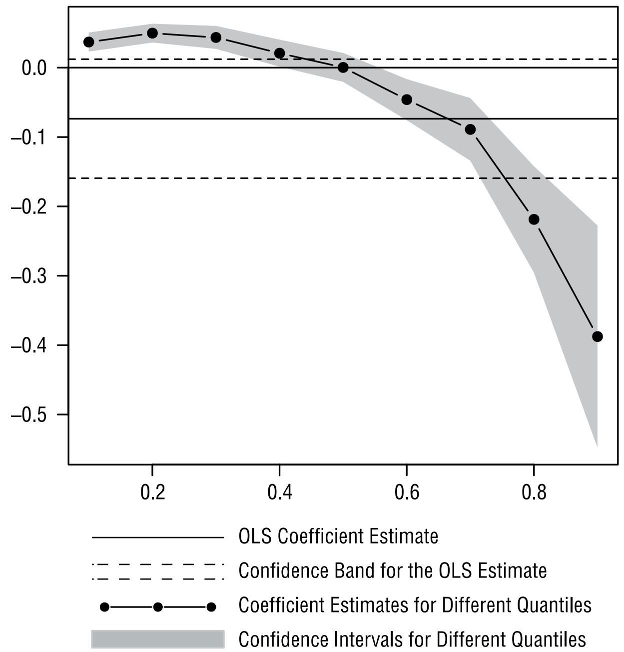

Fig. 2c. **Beta Coefficients of EG Regressor** (Y-Axis) **for Different Quantiles** (X-Axis) **of the Manufacturing Industry in the Model with Net Profit Margin Dependent Variable**

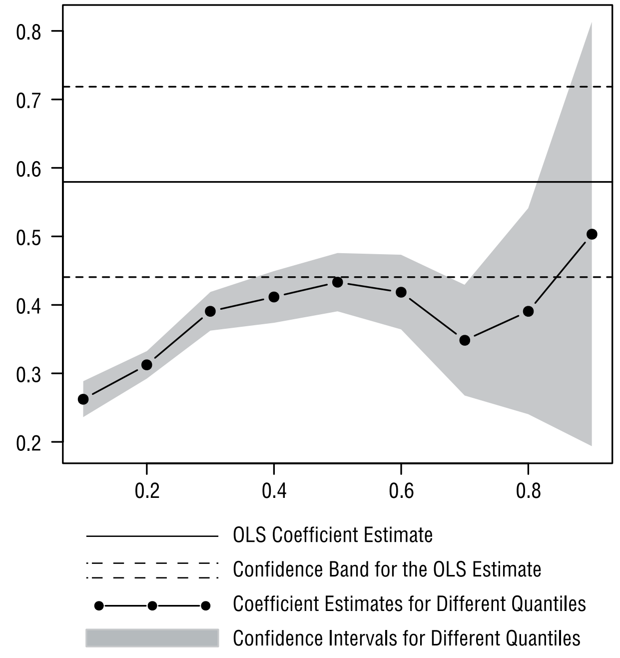

Fig. 2d. **Beta Coefficients of EG Regressor** (Y-Axis) **for Different Quantiles** (X-Axis) **of the Wholesale, Retail and Food Services Industry in the Model with Net Profit Margin Dependent Variable**

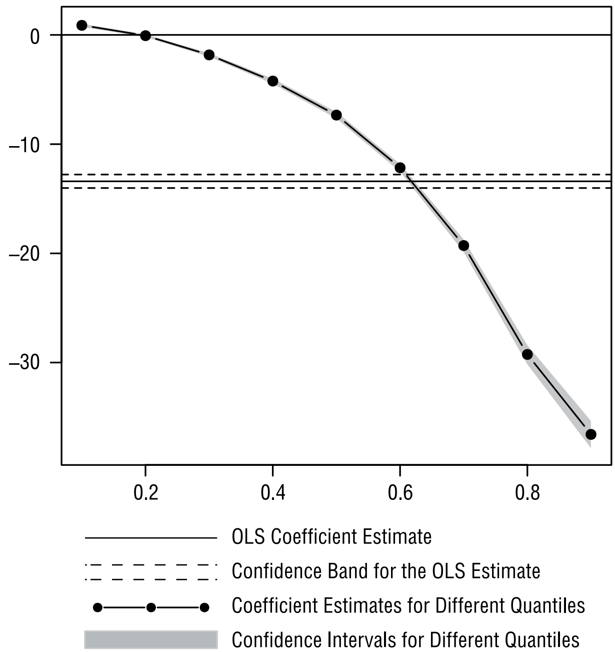

Fig. 2e. **Beta Coefficients of EG Regressor** (Y-Axis) **for Different Quantiles** (X-Axis) **of the Construction Industry in the Model with Net Profit Margin Dependent Variable**

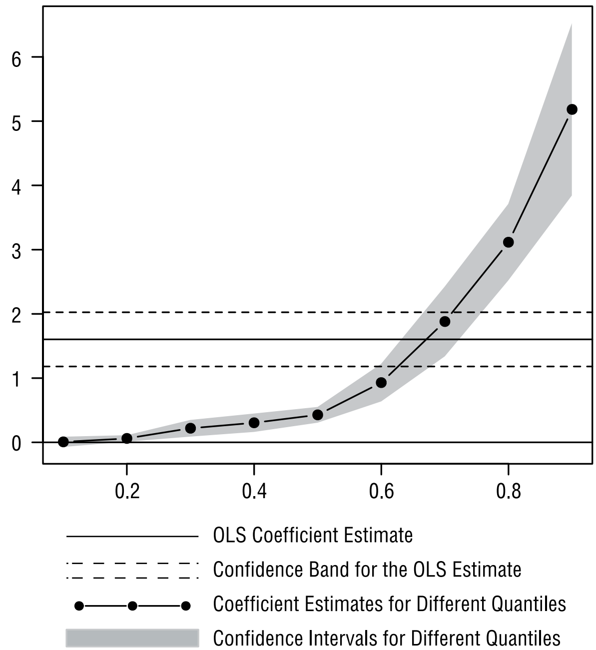

Fig. 2f. **Beta Coefficients of EG Regressor** (Y-Axis) **for Different Quantiles** (X-Axis) **of the Transport Industry in the Model with Net Profit Margin Dependent Variable**

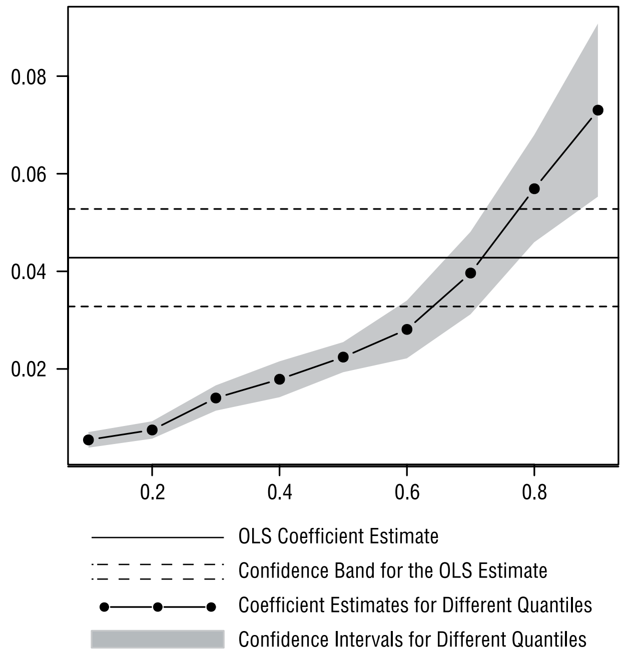

Fig. 2g. **Beta Coefficients of EG Regressor** (Y-Axis) **for Different Quantiles** (X-Axis) **of the IT, Broadcasting, and Telecommunication Industry in the Model with Net Profit Margin Dependent Variable**

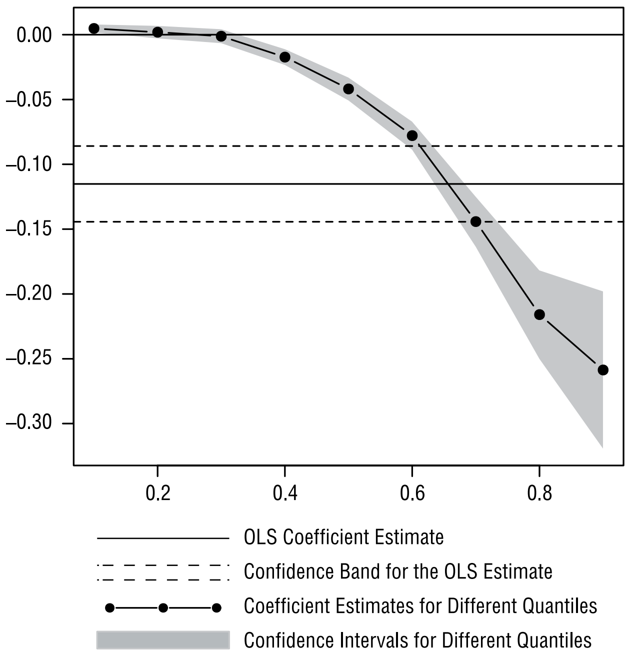

Fig. 2h. **Beta Coefficients of EG Regressor** (Y-Axis) **for Different Quantiles** (X-Axis) **of the Services and Minor Industries Industry in the Model with Net Profit Margin Dependent Variable**

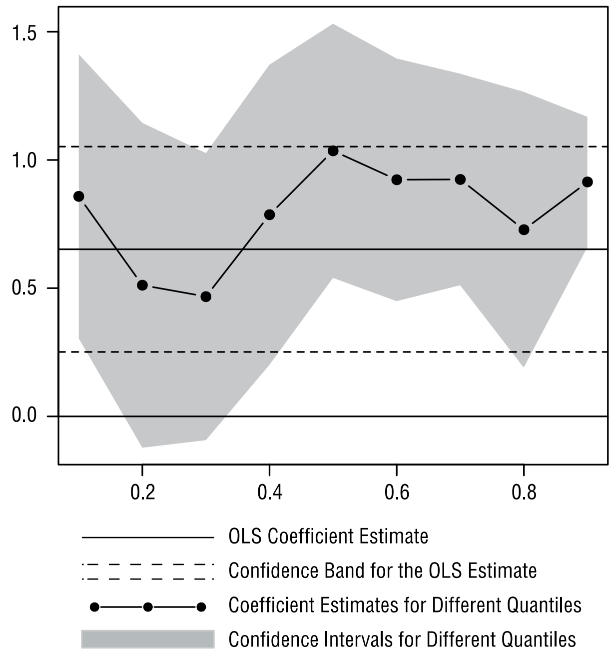

Fig. 3a. **Beta Coefficients of HHI Regressor** (Y-Axis) **for Different Quantiles** (X-Axis) **of the Agriculture, Fishing, and Forestry Industry in the Model with Sales Margin Dependent Variable**

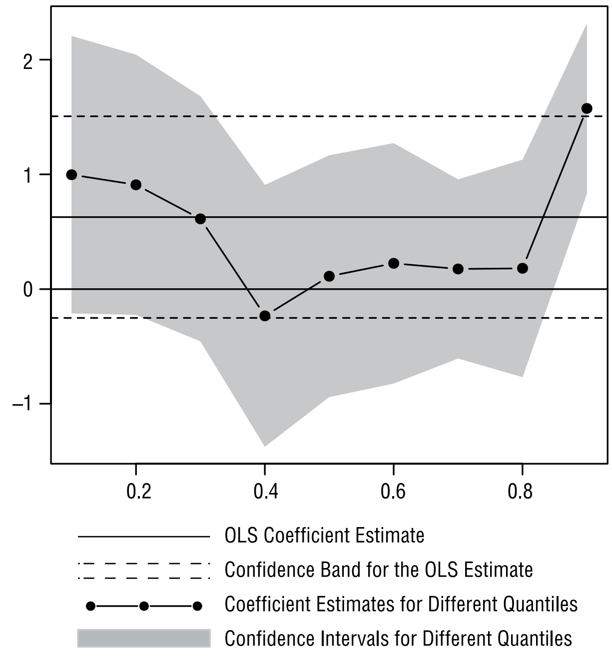

Fig. 3b. **Beta Coefficients of HHI Regressor** (Y-Axis) **for Different Quantiles** (X-Axis) **of the Mining Industry in the Model with Sales Margin Dependent Variable**

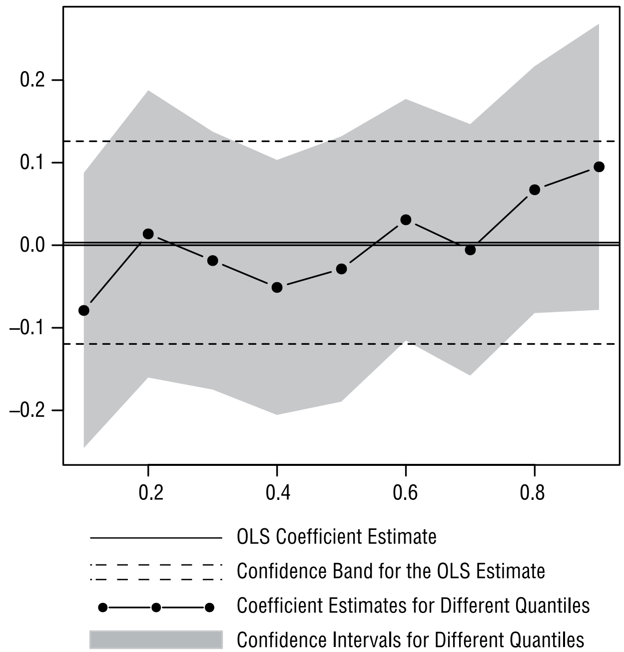

Fig. 3c. **Beta Coefficients of HHI Regressor** (Y-Axis) **for Different Quantiles** (X-Axis) **of the Manufacturing Industry in the Model with Sales Margin Dependent Variable**

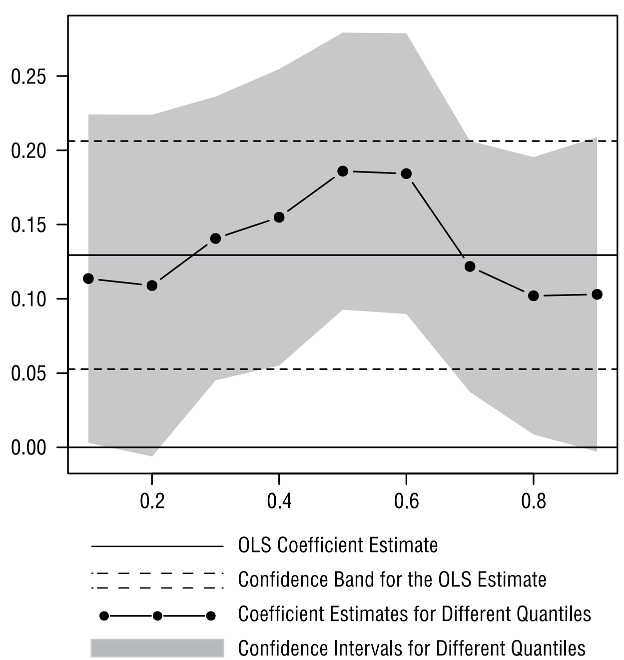

Fig. 3d. **Beta Coefficients of HHI Regressor** (Y-Axis) **for Different Quantiles** (X-Axis) **of the Wholesale, Retail and Food Services Industry in the Model with Sales Margin Dependent Variable**

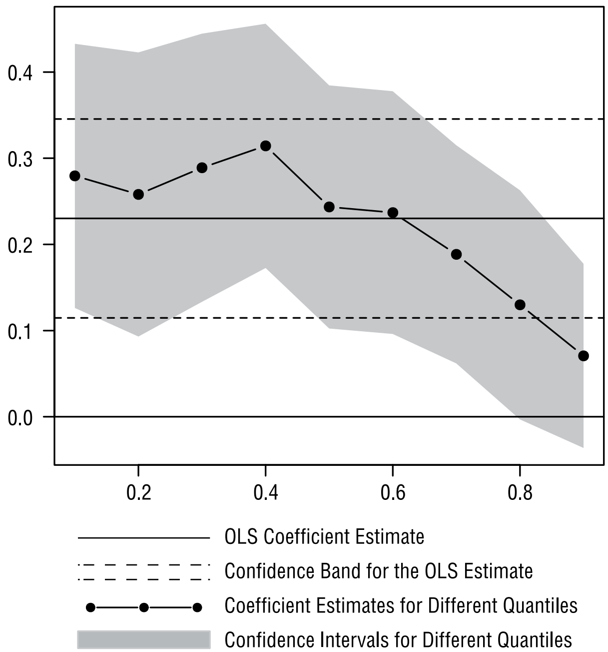

Fig. 3e. **Beta Coefficients of HHI Regressor** (Y-Axis) **for Different Quantiles** (X-Axis) **of the Construction Industry in the Model with Sales Margin Dependent Variable**

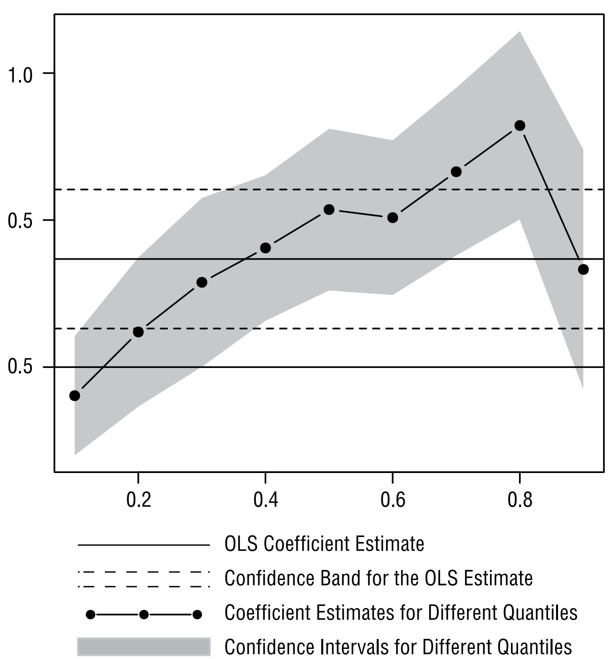

Fig. 3f. **Beta Coefficients of HHI Regressor** (Y-Axis) **for Different Quantiles** (X-Axis) **of the Transport Industry in the Model with Sales Margin Dependent Variable**

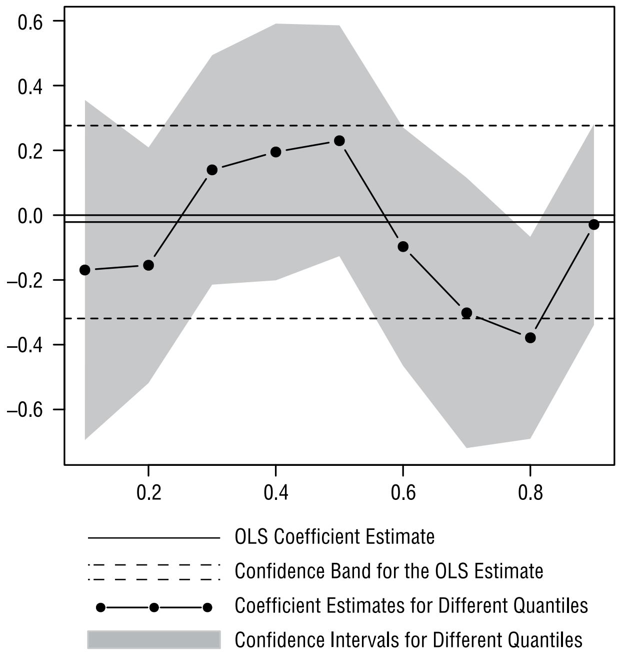

Fig. 3g. **Beta Coefficients of HHI Regressor** (Y-Axis) **for Different Quantiles** (X-Axis) **of the IT, Broadcasting, and Telecommunication Industry in the Model with Sales Margin Dependent Variable**

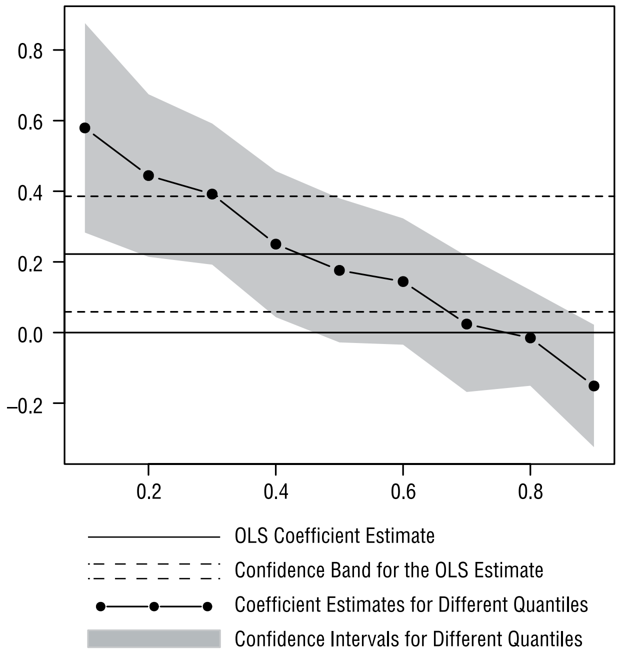

Fig. 3h. **Beta Coefficients of HHI Regressor** (Y-Axis) **for Different Quantiles** (X-Axis) **of the Services and Minor Industries Industry in the Model with Sales Margin Dependent Variable**

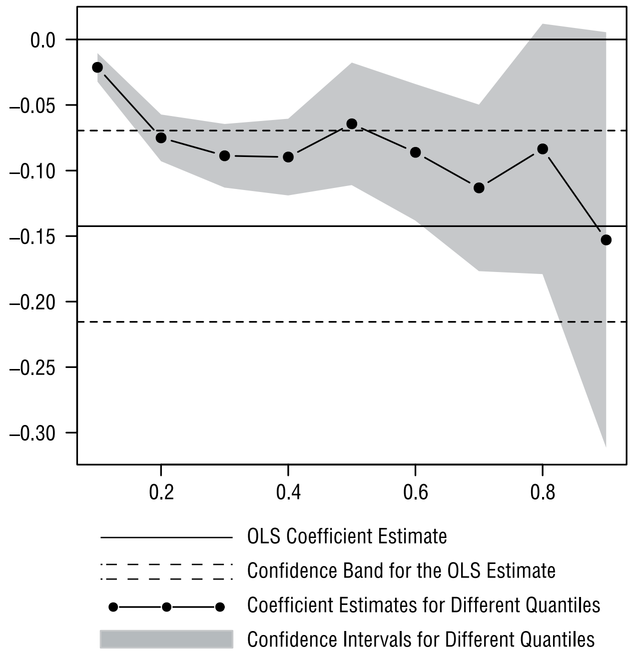

Fig. 4a. **Beta Coefficients of HHI Regressor** (Y-Axis) **for Different Quantiles** (X-Axis) **of the Agriculture, Fishing, and Forestry Industry in the Model with Net Profit Margin Dependent Variable**

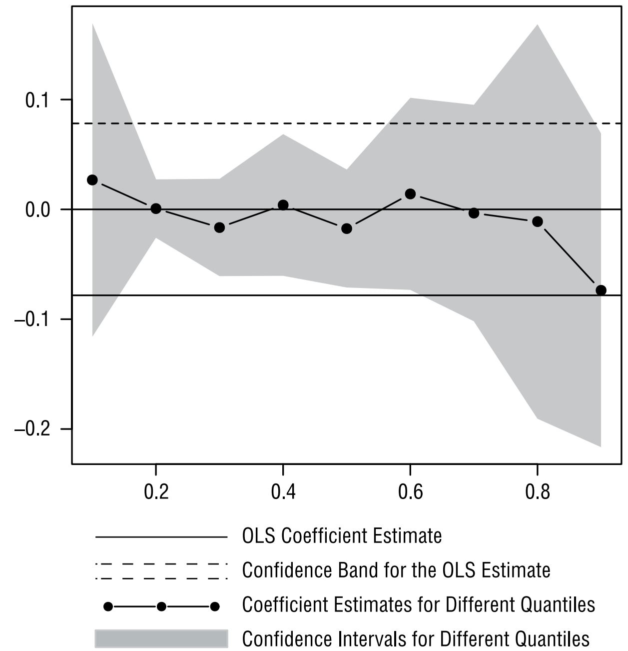

Fig. 4b. **Beta Coefficients of HHI Regressor** (Y-Axis) **for Different Quantiles** (X-Axis) **of the Mining Industry in the Model with Net Profit Margin Dependent Variable**

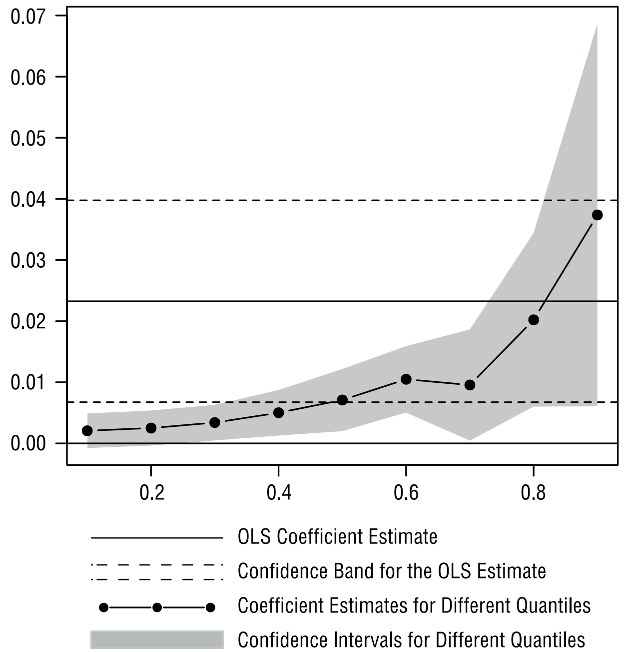

Fig. 4c. **Beta Coefficients of HHI Regressor** (Y-Axis) **for Different Quantiles** (X-Axis) **of the Manufacturing Industry in the Model with Net Profit Margin Dependent Variable**

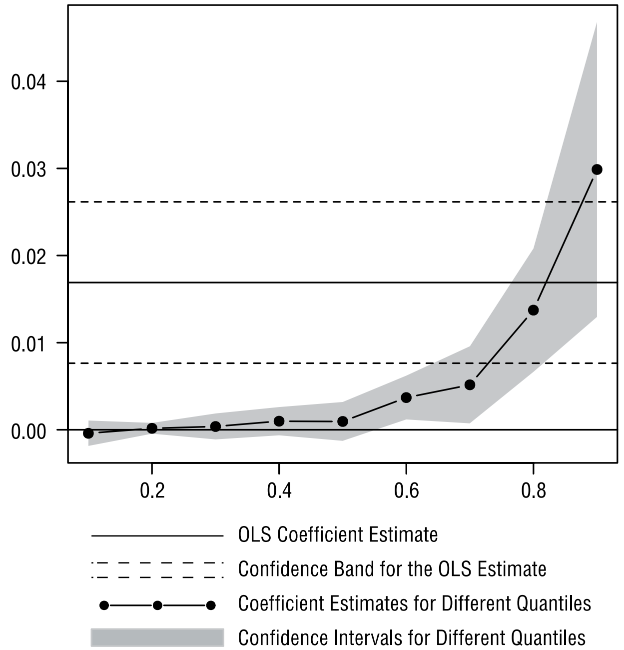

Fig. 4d. **Beta Coefficients of HHI Regressor** (Y-Axis) **for Different Quantiles** (X-Axis) **of the Wholesale, Retail and Food Services Industry in the Model with Net Profit Margin Dependent Variable**

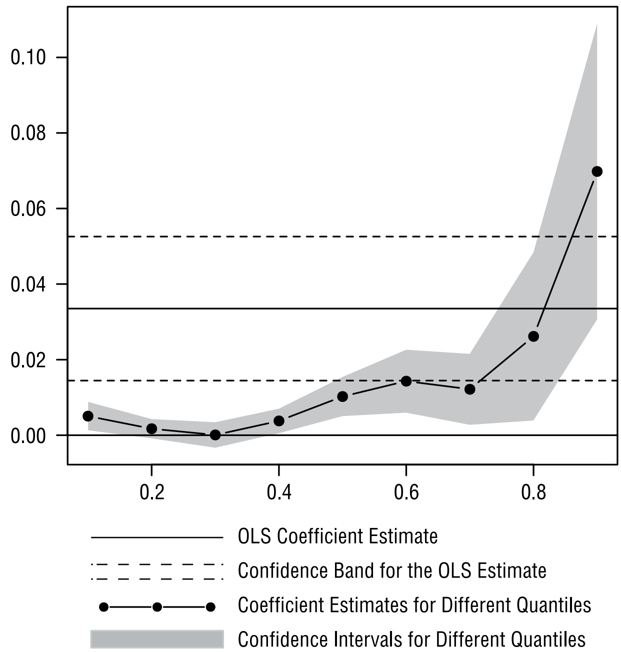

Fig. 4e. **Beta Coefficients of HHI regressor** (Y-Axis) **for Different Quantiles** (X-Axis) **of the Construction Industry in the Model with Net Profit Margin Dependent Variable**

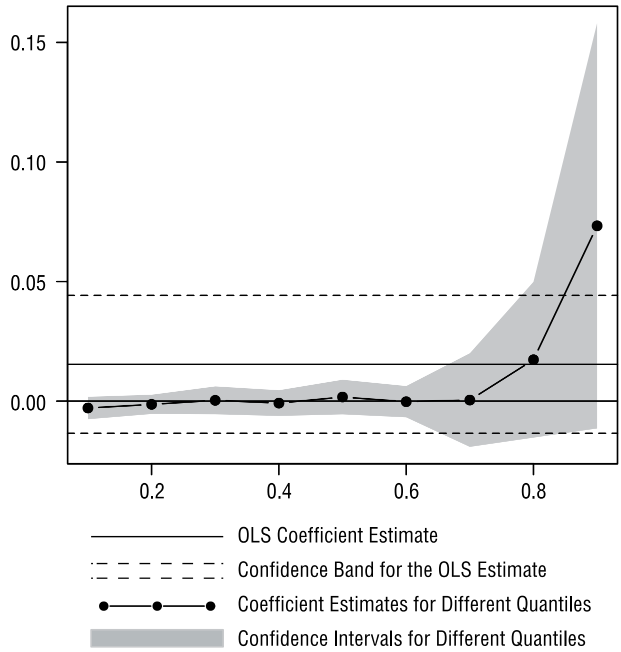

Fig. 4f. **Beta Coefficients of HHI Regressor** (Y-Axis) **for Different Quantiles** (X-Axis) **of the Transport Industry in the Model with Net Profit Margin Dependent Variable**

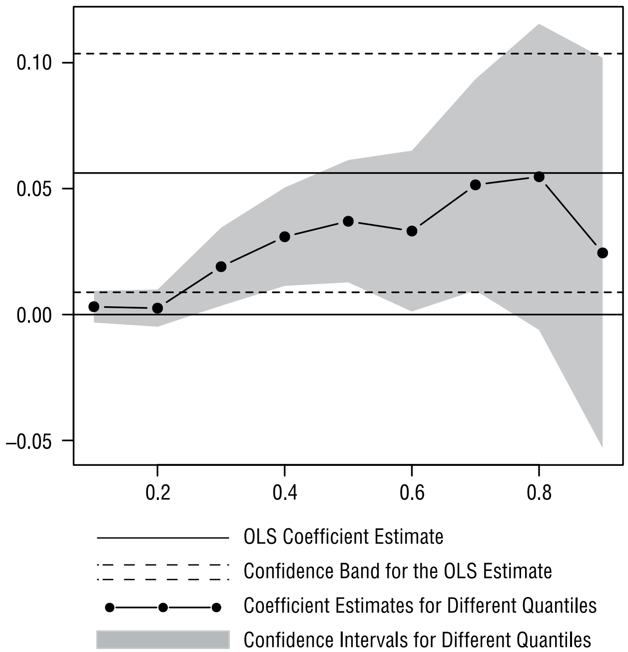

Fig. 4g. **Beta Coefficients of HHI Regressor** (Y-Axis) **for Different Quantiles** (X-Axis) **of the IT, Broadcasting, and Telecommunication Industry in the Model with Net Profit Margin Dependent Variable**

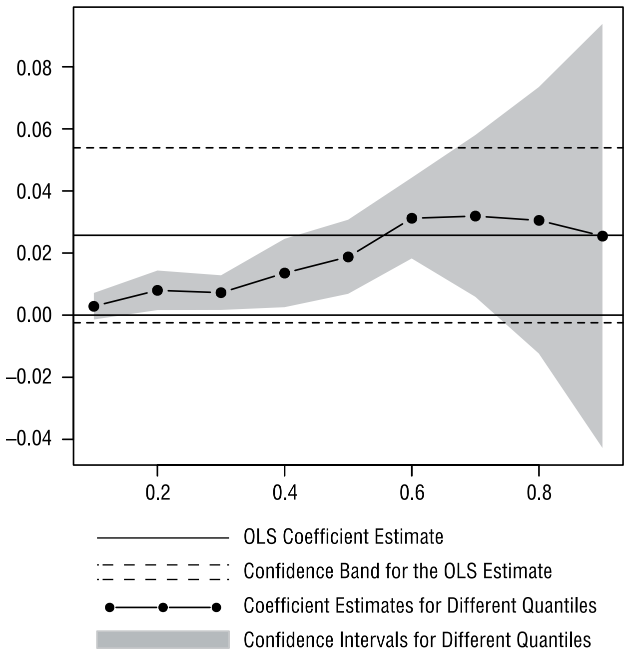

Fig. 4h. **Beta Coefficients of HHI Regressor** (Y-Axis) **for Different Quantiles** (X-Axis) **of the Services and Minor Industries Industry in the Model with Net Profit Margin Dependent Variable**

6. Conclusion

This paper sheds light on the question of how industry localization and regional economy diversification influence the performance of enterprises that operate under different economic activities. Regression analysis has shown that there is strong evidence in support of the idea that localization and diversification effects are notable for companies' performance. It has also been discovered that the effects vary for different industries and for different implied performance within an industry.

Therefore, the two main hypotheses tested in this paper cannot be rejected for the majority of industries. The only exception is the *Mining* industry, where neither localization nor diversification effects are somehow connected to the companies' business efficiency. The reason is that *Mining* companies' locations are usually straightly tied to the resource fields and their performance is not connected to agglomeration.

The negative influence of the degree of regional economy diversification on firm performance has been found for almost all the industries (except *Agriculture, Fishing, and Forestry* in terms of net profit margin). These results are in line with the theoretical findings stated in Jacobs's book [Jacobs, 1961], which stated that the influence of regional diversification has to be restricted due to the fact that regional markets are bound and firms import their goods and services. Industries where the influence has been found (*Wholesale, Retail and Food Services, Construction* and *Other Services*) are specific in the way that the enterprises usually work within their home market, and only large-scale firms take their businesses beyond the regional borders.

The benefits that localization brings are well-known: access to the market on both supply and demand sides, lower costs of hiring employees, lower transportation costs, and knowledge spillover effects [Davidson, Mariev, 2015; Holmes, 1999; Rosenthal, Strange, 2006]. However, for some industries, considering their peculiarities and industry specific features, these positive externalities were overweighed by the negative ones—such as the more expensive ground rent payments for the *Construction* industry and the higher rental payments for *Service* companies, or even intense competition within the local territories [Lu et al., 2012] (for *Service* companies as well).

Quantile regression analysis has made it possible to discover that localization and diversification effects are different for high and low margin companies even within an industry. If a company is a newcomer or is less advanced that its rivals, which results in comparatively lower performance rates, then it is likely that such a company would benefit more greatly from an increasing degree of localization. Higher margin companies, on the contrary, are more likely to appear as donors rather than recipients. Knowledge and technology spillovers and easy adoption of the best business models are the channels to hold back strong competitors and support developing companies [Audretsch, Feldman, 1996; Shaver, Flyer, 2000].

Localization and diversification effects can also be measured quantitatively. Every 0.1 change in the EG index (localization) results in a 1%–4% change in the sales margin and a 1–1,2% change in the net profit margin depending on the industry in absolute terms. Diversification's contribution in firms' performance is lower: every 0.1 change in the HHI changes companies' sales margin by 1%–1,5% and net profit margin by up to 1–1.05%. An exception from this is the *Construction* industry, where stronger (up to 7-fold drop in sales margin and 5% drop in net profit margin) negative effects have been found.

Models and findings elaborated on in the present paper can be useful when formulating regional policy. The results can also be applied to regulatory purposes in setting up stimulation measures for some industries. Finally, the findings are a good basis for constructing models of optimal allocation of industries inside the agglomeration or that of industries across the country.

There are a few important limitations of the research that need to be discussed. First, it is based on cross-section data, so no dynamics can be measured, and time consistency cannot be checked. Secondly, the present paper does not contain formal causal checks, relating the obtained effects to the existing literature. Finally, the methodology is simplified due to the content of the available data—however, the study of localization and diversification effects might be continued if the analysis includes the presence, weight, and spatial distribution of big cities in a region, which may possibly especially strengthen the diversification effects.

References

1. Aleksandrova E., Behrens K., Kuznetsova M. Manufacturing (Co)Agglomeration in a Transition Country: Evidence from Russia. *Journal of Regional Science*, 2020, vol. 60, no. 1, pp. 88-128. DOI:10.1111/jors.12436.
2. Audretsch D. B. Agglomeration and the Location of Innovative Activity. *Oxford Review of Economic Policy*, 1998, vol. 14, no. 2, pp. 18-29.
3. Audretsch D. B., Feldman M. R&D Spillovers and the Geography of Innovation and Production. *American Economic Review*, 1996, vol. 86, no. 3, pp. 630-640.
4. Baptista R., Swann P. Do Firms in Clusters Innovate More? *Research Policy*, 1998, vol. 27, no. 5, pp. 525-540.
5. Baumgardner J. R. Physicians' Services and the Division of Labor Across Local Markets. *Journal of Political Economy*, 1988, vol. 96, no. 5, pp. 948-982.
6. Cassey A. J., Smith B. O. Simulating Confidence for the Ellison-Glaeser Index. *Journal of Urban Economics*, 2014, vol. 81, pp. 85-103. DOI:10.1016/j.jue.2014.02.005.
7. Ciccone A., Hall R. E. Productivity and the Density of Economic Activity. *American Economic Review*, 1996, vol. 86, no. 1, pp. 54-70.

8. Cohen J. P., Morrison Paul C. J. Agglomeration, Productivity and Regional Growth: Production Theory Approaches. In: *Handbook of Regional Growth and Development Theories*, Capello R., Nijkamp P. (eds.). Cheltenham, Edward Elgar, 2009, pp. 101-117.
9. Davidson N. B., Mariev O. S. The Impact of Spatial Concentration on Enterprise Performance. *Economy of Region*, 2015, no. 4, pp. 95-105. DOI:10.17059/2015-4-8.
10. Dumais G., Ellison G., Glaeser E. L. Geographic Concentration as a Dynamic Process. *The Review of Economics and Statistics*, 1997, vol. 84, no. 2, pp. 193-204.
11. Duranton G., Puga D. Micro-Foundations of Urban Agglomeration Economies. In: *Handbook of Regional and Urban Economics: Cities and Geography,* Henderson J. V., Thisse J. F. (eds.). North Holland, Elsevier Inc., 2004, vol. 4, pp. 2063-2117.
12. Ellison G., Glaeser E. L. Geographic Concentration in U.S. Manufacturing Industries: A Dartboard Approach. *Journal of Political Economy*, 1997, vol. 105, no. 5, pp. 889-927.
13. Henderson J. V. Marshall's Scale Economies. *Journal of Urban Economics*, 2003, vol. 53. no. 1, pp. 1-28.
14. Hervas-Oliver J. L., Sempere-Ripoll F., Rojas Alvarado R., Estelles-Miguel S. Agglomerations and Firm Performance: Who Benefits and How Much? *Regional Studies*, 2018, vol. 52, no. 3, pp. 338-349. DOI:10.1080/00343404.2017.1297895.
15. Holmes T. J. Localization of Industry and Vertical Disintegration. *The Review of Economics and Statistics*, 1999, vol. 81, no. 2, pp. 314-325.
16. Jacobs J. *The Death and Life of Great American Cities.* New York, NY, Random House, 1961.
17. Kolomak E. A. Evolution of Spatial Distribution of Economic Activity in Russia. *Regional Research of Russia*, 2015a, vol. 5, no. 3, pp. 236-242. DOI:10.1134/S2079970515030065.
18. Kolomak E. A. Spatial Externalities as a Source of Economic Growth. *Region: Economics and Sociology,* 2010, no. 4, pp. 73-87. (In Russ.)
19. Kolomak E. A. Urbanization Resource in Russia. *Prostranstvennaya ekonomika [Spatial Economics]*, 2015b, no. 4, pp. 59-74. (In Russ.)
20. Lavrinenko P. A., Mikhailova T. N., Romashina A. A., Chistyakov P. A. Agglomeration Effect as a Tool of Regional Development. *Problemy prognozirovaniya [Studies on Russian Economic Development]*, 2019, vol. 30, no. 3, pp. 50-59. (In Russ.)
21. Loecker J. De, Warzynski F. Markups and Firm-Level Export Status. *American Economic Review*, 2012, vol. 102, no. 6, pp. 2437-2471. DOI:10.1257/aer.102.6.2437.
22. Lu Y., Tao Z., Yu L. *Agglomeration and Markup.* University Library of Munich, MPRA Paper, no. 38974, 2012.
23. Malmberg A., Malmberg B., Lundequist P. Agglomeration and Firm Performance: Economies of Scale, Localisation, and Urbanisation Among Swedish Export Firms. *Environment and Planning A: Economy and Space*, 2000, vol. 32, no. 2, pp. 305-321.
24. Marshall A. *Principles of Economics.* L., Macmillan, 1920.
25. Martin P., Mayer T., Mayneris F. Spatial Concentration and Plant-Level Productivity in France. *Journal of Urban Economics*, 2011, vol. 69, no. 2, pp. 182-195.
26. Mikhailova T. N. Transformation of the Industrial Geography of the Russian Federation. *Rossiyskoe predprinimatelstvo [Russian Journal of Entrepreneurship]*, 2016, vol. 17, no. 3, pp. 351-358. DOI:10.18334/rp.17.3.2223. (In Russ.)
27. Morrison Paul C. J., Siegel D. S. Knowledge Capital and Cost Structure in the U.S. Food and Fiber Industries. *American Journal of Agricultural Economics*, 1998, vol. 80, no. 1, pp. 30-45.
28. Nakamura R. Agglomeration Economies in Urban Manufacturing Industries: A Case of Japanese Cities. *Journal of Urban Economics*, 1985, vol. 17, no. 1, pp. 108-124.
29. Quigley J. M. Urban Diversity and Economic Growth. *Journal of Economic Perspectives*, 1998, vol. 12, no. 2, pp. 127-138.
30. Rosenthal S. S., Strange W. C. Evidence on the Nature and Sources of Agglomeration Economies. In: *Handbook of Regional and Urban Economics: Cities and Geography*, Henderson J. V., Thisse J. (eds.). North Holland, Elsevier Inc., vol. 4, 2004, pp. 2119-2171.
31. Rosenthal S. S., Strange W. C. The Micro-Empirics of Agglomeration Economies. In: *A Companion to Urban Economics,* Arnott R. J., McMillen D. P. (eds.). Hoboken, NJ, Wiley-Blackwell, 2006. DOI:10.1002/9780470996225.

32. Shaver J. M., Flyer F. Agglomeration Economies, Firm Heterogeneity, and Foreign Direct Investment in the United States. *Strategic Management Journal*, 2000, vol. 21, no. 12, pp. 1175-1193.
33. Vorobyev P. V., Kislyak N. V., Davidson N. B. Spatial Concentration and Firm Performance in Russia. *EERC Working Paper Series,* no. 10/05e, 2010.
34. Wheaton W. C., Lewis M. J. Urban Wages and Labor Market Agglomeration. *Journal of Urban Economics*, 2002, vol. 51, no. 3, pp. 542-562.
35. Zyuzin A. V., Demidova O. A., Dolgopyatova T. G. Localization and Diversification of Russian Economy: Regions' and Industries' Peculiarities. *Prostranstvennaya ekonomika [Spatial Economics]*, 2020, vol. 16, no. 2, pp. 39-69. DOI:10.14530/se.2020.2.039-069. (In Russ.)